PCT





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

A61K

(11) International Publication Number: WO 99/60986

(43) International Publication Date: 2 December 1999 (02.12.99)

US

(21) International Application Number: PCT/US99/11743

(22) International Filing Date: 27 May 1999 (27.05.99)

(30) Priority Data: 09/085,199 27 May 1998 (27.05.98)

(71) Applicants (for all designated States except US): UNIVER-SITY OF BRITISH COLUMBIA [CA/CA]; University Industry Liaison Office, IRC Building – Room 331, 2194 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3 (CA). MERCK FROSST CANADA & CO. [CA/CA]; PO/CP 1005, Pointe Claire-Dorval, Québec H9R 4P8 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): KALCHMAN, Michael [CA/CA]; #1403–900 Yonge Street, Toronto, Ontario M4W 3P5 (CA). HAYDEN, Michael, R. [US/CA]; 4484 West Seventh, Vancouver, British Columbia V6R 1W9 (CA). HACKAM, Abigail [CA/CA]; 1420 West 11th Avenue, Vancouver, British Columbia V6H 1L2 (CA). CHOPRA, Vikramjit [CA/CA]; Suite 210, 2475 Blenheim Street, Vancouver, British Columbia V6K 4N7 (CA). NICHOLSON, Donald, W. [CA/CA]; 18–750 Milton Street, Montréal, Québec H2X 1W4 (CA). VALLAINCOURT, John, P. [CA/CA]; 18022 Amalfi Street, Québec, Québec H9K 1N7

(CA). RASPER, Dita, M. [CA/CA]; Apartment #7, 16203 Pierrefonds Boulevard, Pierrefonds, Québec H9H 4S8 (CA).

(74) Agent: COPPOLA, Joseph, A.; Merck & Co., Inc., P.O. Box 2000, Rahway, NJ 07065-0907 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: APOPTOSIS MODULATORS THAT INTERACT WITH THE HUNTINGTON'S DISEASE GENE

(57) Abstract

A family of proteins, including a specific human protein designated as HIP1, has been identified that interact differently with the gene product of a normal (16 CAG repeat) and an expanded (>44 CAG repeat) HD gene. Expression of the HIP1 protein was found to be enriched in the brain. Analysis of the sequence of the HIP1 protein indicated that it includes a death effector domain (DED), suggesting an apoptotic function. Thus, it appears that a normal function of Huntingtin may be to bind HIP1 and related apoptosis modulators, reducing its effectiveness in stimulating cell death. Since expanded huntingtin performs this function less well, there is an increase in HIP1-modulated cell death in individuals with an expanded repeat in the HD gene. This understanding of the likely role of huntingtin and HIP1 or related proteins (collectively "HIP-apoptosis modulating proteins") in the pathology of Huntington's disease offers several possibilities for therapy. First, because the function of huntingtin apparently depends at least in part on the ability to interact with HIP-apoptosis modulating proteins, added expression (e.g., via gene therapy) of normal (non-expanded) huntingtin or of the HIP-binding region of huntingtin should proteins with the death signaling complex. Alternatively, a mutant form of HIP-protein from which the DED has been deleted might be introduced, for example using gene therapy techniques. Because HIP-apoptosis modulating proteins have been shown to self-associate, a protein with a deleted DED may compete with endogenous HIP-protein in the formation of these associations, thereby reducing the amount of apoptotically-active HIP-protein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

	00000 0000 10 1011111				• •		
AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Мопасо	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	ΙT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JР	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugosłavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

WO 99/60986 PCT/US99/11743

APOPTOSIS MODULATORS THAT INTERACT WITH THE HUNTINGTON'S DISEASE GENE

BACKGROUND OF THE INVENTION

5

10

15

20

25

This application relates to a family of apoptosis modulators that interact with the Huntington's Disease gene product, and to methods and compositions relating thereto.

"Interacting proteins" are proteins which associate *in vivo* to form specific complexes. Non-covalent bonds, including hydrogen bonds, hydrophobic interactions and other molecular associations form between the proteins when two protein surfaces are matched or have affinity for each other. This affinity or match is required for the recognition of the two proteins, and the formation of an interaction. Protein-protein interactions are involved in the assembly of enzyme subunits; in antigen-antibody reactions; in forming the supramolecular structures of ribosomes, filaments, and viruses; in transport; and in the interaction of receptors on a cell with growth factors and hormones.

Huntington's disease is an adult onset disorder characterized by selective neuronal loss in discrete regions of the brain and spinal chord that lead to progressive movement disorder, personality change and intellectual decline. From onset, which generally occurs around age 40, the disease progresses with worsening symptoms, ending in death approximately 18 years after onset.

The biochemical cause of Huntington's disease is unclear. While the biochemical cause of Huntington's disease has remained elusive, a mutation in a gene within chromosome 4p16.3 subband has been identified and linked to the disease. This gene, referred to as the Huntington's Disease or HD gene, contains two repeat regions, a CAG repeat region and a CCG repeat region. Testing of Huntington's disease patients has shown that the CAG region is highly polymorphic, and that the number of CAG repeat units in the CAG repeat region is a very reliable indicator of having inherited the gene for Huntington's disease. Thus, in control individuals and in most individuals suffering from neuropsychiatric disorders other than Huntington's disease, the number of CAG repeats is between 9 and 35, while in individuals suffering from Huntington's disease the number of CAG repeats is expanded and is 36 or greater.



To date, no differences have been observed at either the total RNA, mRNA or protein levels between normal and HD-affected individuals. Thus, the function of the HD protein and its role in the pathogenesis of Huntington's Disease remain to be elucidated.

SUMMARY OF THE INVENTION

5

10

15

20

25

30

We have now identified a protein designated as HIP1, that interact differently with the gene product of a normal (16 CAG repeat) and an expanded (>44 CAG repeat) HD gene. The HIP1 protein originally isolated from a yeast two-hybrid screen is encoded by a 1.2 kb cDNA (Seq. ID. No. 1), devoid of stop codons, that is expressed as a 400 amino acid polypeptide (Seq. ID. No. 2). Subsequent study has elucidated additional sequence for HIP1 such that a 1090 amino acid protein is now known. (Seq. ID No. 5). Expression of the HIP1 protein was found to be enriched in the brain.

Analysis of the sequence of the HIP1 protein indicated that it includes a death effector domain (DED), suggesting an apoptotic function. Thus, it appears that a normal function of huntingtin may be to bind HIP1 and related apoptosis modulators, reducing its effectiveness in stimulating cell death. Since expanded huntingtin performs this function less well, there is an increase in HIP1-modulated cell death in individuals with an expanded repeat in the HD gene. Furthermore, additional members of the same family of proteins have been identified which also contain a DED. Thus, the present invention provides a new class of apoptotic modulators which are referred to as HIP-apoptosis modulating proteins.

This understanding of the likely role of huntingtin and HIP1 or related proteins in the pathology of Huntington's Disease offers several possibilities for therapy. First, because the function of huntingtin apparently depends at least in part on the ability to interact with HIP-apoptosis modulating proteins, added expression (e.g., via gene therapy) of normal (non-expanded) huntingtin or of the HIP-binding region of huntingtin should provide a therapeutic benefit. Other DED-interacting peptides could also be used to mask and reduce the interaction of HIP-apoptosis modulating proteins with the death signaling complex. Alternatively, a mutant form of HIP-protein from which the DED has been deleted might be introduced, for example using gene therapy techniques. Because HIP-apoptosis modulating proteins have been shown to self-associate, a protein with a deleted DED may compete with

10

15

20

25

30

endogenous HIP-protein in the formation of these associations, thereby reducing the amount of apoptotically-active HIP-protein.

BRIEF DESCRIPTION OF THE DRAWING

- Fig. 1 graphically depicts the amount of interaction between HIP1 and Huntingtin proteins with varying lengths of polyglutamine repeat;
 - Fig. 2 compares the nucleic acid sequences of human and murine HIP1 and HIP1a;
 - Fig. 3 compares the amino acid sequences of human and murine HIP1 and HIP1a;
- Fig. 4 shows the sequences of various death effector domains in comparison to the DED of human and murine HIP1 and HIP1a;
 - Fig. 5 shows the genomic organization of human HIP1;
 - Fig. 6 compares the sequences of human HIP1 with ZK370.3 protein of *C. elegans*;
- Fig. 7 shows mouse EST's with homology to human HIP1 cDNA used to screen a mouse brain library;
 - Fig. 8 shows the affect of HIP1 on susceptibility of cells to stress; and
- Figs. 9A 9C show the toxicity of HIP1 in the presence of huntingtin with different lengths of polyglutamine repeats.

DETAILED DESCRIPTION OF THE INVENTION

This application relates to a new family of proteins function as modulators of apoptosis. At least some of these proteins, notably the human protein designated HIP1, interact with the gene product of the Huntington's disease gene. Other proteins within the family possess at least 40% and preferably more than 50% nucleotide identity with HIP1 and include a death effector domain (DED). Such proteins are referred to in the specification and claims hereof as "HIP-apoptosis modulating proteins."

The first HIP-apoptosis modulating protein identified was designated as HIP1. HIP1 was identified using the yeast two-hybrid system described in US Patent No. 5,283,173 which is incorporated herein by reference. Briefly, this system utilizes two chimeric genes or plasmids expressible in a yeast host. The yeast host is selected to contain a detectable marker gene having a binding site for the DNA binding domain of a transcriptional activator. The



10

15

20

25

30

first chimeric gene or plasmid encodes a DNA-binding domain which recognizes the binding site of the selectable marker gene and a test protein or protein fragment. The second chimeric gene or plasmid encodes for a second test protein and a transcriptional activation domain. The two chimeric genes or plasmids are introduced into the host cell and expressed, and the cells are cultivated. Expression of the detectable marker gene only occurs when the gene product of the first chimeric gene or plasmid binds to the DNA binding domain of the detectable marker gene, and a transcriptional activation domain is brought into sufficient proximity to the DNA-binding domain, an occurrence which is facilitated by protein-protein interactions between the first and second test proteins. By selecting for cells expressing the detectable marker gene, those cells which contain chimeric genes or plasmids for interacting proteins can be identified, and the gene can be recovered and identified.

In testing for Huntington Interacting Proteins, several different plasmids were prepared containing portions of the human HD gene. The first four, identified as 16pGBT9, 44pGBT9, 80pGBT9 and 128pGBT9, were GAL4 DNA binding domain-HD in-frame fusions containing nucleotides 314 to 1955 (amino acids 1-540) of the published HD cDNA sequences cloned into the vector pGBT9 (Clontech). These plasmids contain a CAG repeat region of 16, 44, 80 and 128 glutamine-encoding repeats, respectively. A clone (DMK BamHIpGBT9) was made by fusing a cDNA encoding the first 544 amino acids of the myotonic dystrophy gene (a gift from R. Korneluk) in-frame with the GAL4-DNA BD of pGBT9 and was used as a negative control.

These plasmids have been used to identify and characterize HIP1, as well as two additional HD-interacting proteins, HIP2 and HIP3, which have not yet been tested for function as apoptosis modulators. These plasmids can be further used for the identification of additional interacting proteins which do act as apoptosis modulators, and for tests to refine the region on the protein in which the interaction occurs. Thus, one aspect of the invention is these four plasmids, and the use of these plasmids in identifying HD-interacting proteins. Furthermore, it will be appreciated that the GAL4 DNA-binding and activating domains are not the only domains which can be used in the yeast two-hybrid assay. Thus, in a broader sense, the invention encompasses any chimeric genes or plasmids containing nucleotides 314 to 1955 of the HD gene together with an activating or DNA-binding domain suitable for use

10

15

20

25

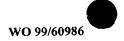
30

in the yeast one, two- or three-hybrid assay for proteins critical in either binding to the HD protein or responsible for regulated expression of the HD gene.

After introducing the plasmids into Y190 yeast host cells, transforming the host cells with an adult human brain Matchmaker™ (Clontech) cDNA library coupled with a GAL4 activating domain, and selecting for the expression of two detectable marker genes to identify clones containing genes for interacting proteins, the activating domain plasmids were recovered and analyzed. As a result of this analysis, three different cDNA fragments were identified as encoding for HD-interacting proteins and designated as HIP1, HIP2 and HIP3. The nucleic acid sequence of HIP1, as originally recovered in the yeast two-hybrid assay, is given in Seq. ID. No 1. The polypeptide which it encodes is given by Seq. ID No. 2. Further investigation of the HIP1 cDNA resulted in the characterization of a longer region of cDNA totaling 4795 bases and a corresponding protein, the sequences of which are given by Seq ID Nos. 3 and 4, respectively. A further portion of the HIP1 protein was characterized, extending the length to the complete protein sequence of 1090 amino acids (Seq. ID No. 5)

The cDNA molecules encoding HIP-apoptosis modulating proteins, particularly those encoding portions of HIP1, can be explored using oligonucleotide probes for example for amplification and sequencing. In addition, oligonucleotide probes complementary to the cDNA can be used as diagnostic probes to localize and quantify the presence of HIP1 DNA. Probes of this type with a one or two base mismatch can also be used in site-directed mutagenesis to introduce variations into the HIP1 sequence which may increase or decrease the apoptotic activity. Preferred targets for such mutations would be the death effector domains. Thus, a further aspect of the present invention is an oligonucleotide probe, preferably having a length of from 15-40 bases which specifically and selectively hybridizes with the cDNA given by Seq. ID No. 1 or 3 or a sequence complementary thereto. As used herein, the phrase "specifically and selectively hybridizes with" the cDNA refers to primers which will hybridize with the cDNA under stringent hybridization conditions.

Probes of this type can also be used for diagnostic purposes to characterize risk of Huntington's Disease like symptoms arising in individuals where the symptoms are present in the family history but are not associated with an expansion of the CAG repeat. Such symptoms may arise from a mutation in HIP1 or other HIP-apoptosis modulating protein



10

15

20

25

which alters the interaction of this protein with huntingtin, thereby increasing the apoptotic activity of the protein even in the presence of a normal (non-expanded) huntingtin molecule. An appropriate probe for this purpose would one which hybridizes with or adjacent to the huntingtin binding region of the HIP-apoptosis modulating protein. In HIP1, this lies within amino acids 129-514.

DNA sequencing of the HIP1 cDNA initially isolated from the yeast two-hybrid screen (Seq. ID No. 1) revealed a 1.2 kb cDNA that shows no significant degree of nucleic acid identity with any stretch of DNA using the blastn program at ncbi (blast@ncbi.n1m.nih.gov). When the larger HIP1 cDNA sequence (SEQ ID NO. 3) was translated into a polypeptide, the HIP1 cDNA coding (nucleotides 328-3069) is observed to be devoid of stop codons, and to produce a 914 amino acid polypeptide. A polypeptide identity search revealed an identity match over the entire length of the protein (46% conservation) with that of a hypothetical protein from C. elegans (ZK370.3 protein; C. elegans cosmid ZK370). This C. elegans protein shares identity with the mouse talin gene, which encodes a 217 kDa protein implicated with maintaining integrity of the cytoskeleton. It also shares identity with the SLA2/MOP2/ END4 gene from Saccharomyces cerevisiae, which is known to code for an essential cytoskeletal associated gene required for the accumulation and or maintenance of plasma membrane H⁺- ATPase on the cell surface. When pairwise comparisons are performed between HIP1 and the C. elegans ZK370.3 protein (Genpept accession number celzk370.3), it shows 26% complete identity and an overall 46% level of conservation. Comparative analysis between HIP1 and SLA2/MOP2/ END4 (EMBL accession number Z22811) demonstrate similar conservation (20% identity, 40% conservation).

Further exploration revealed several important facts about HIP1 that implicate it in a significantly in the pathogenesis of Huntington's Disease. First, as shown in Fig. 1, it was found that the native interaction between HD protein and HIP1 is influenced by the number of CAG repeats. Second, it was found that expression of the HIP1 protein is enriched in the brain. The highest amounts of expression are in the cortex, with lower levels being seen in the cerebellum, caudate and putamen.

It has also been observed that huntingtin proteins with expanded polyglutamine tracts can aggregate into large, irregularly shaped deposits in HD brains, transgenic mice and *in vitro* cell culture. We have shown that in HEK (human embryonic kidney) 293T cells, the aggregation of full-length and smaller huntingtin fragments occurs after the cells have been exposed to a period of apoptotic stress. Martindale, et al., *Nature Genetics* 18: 150-154 (1998). In order to assess the consequence of HIP1 expression in cultured cells, we used huntingtin aggregation as one marker of viability. What we found was that cells cotransfected with huntingtin (128 CAG repeats) and HIP1 contained aggregates comparable to those observed following application of apoptotic stress with sub-lethal doses of tamoxifen in 14% of the cells, and that these cells were the ones in which both genes had been introduced as reflected by a double marker experiment. Transfection of a gene encoding a fusion protein of 128 repeat huntingtin and the DED domain from HIP1 ligated in the sense orientation resulted in aggregate formation in 30 to 50% of the cells.

5

10

15

20

25

30

The implications of the apoptotic activity of HIP1 are two-fold. First, the fact that this activity is apparently differentially modulated by interaction with huntingtin having normal and expanded repeats implicates HIP1 in the apoptotic neuronal death which is observed in Huntington's disease and makes HIP1 a logical target for therapy. A second implication of the apoptotic activity of HIP1 is the potential for use of HIP1 as a therapeutic agent to introduce apoptosis in cancer cells.

Therapeutic targeting of HIP1 or other HIP-apoptosis modulating proteins might take any of several forms, but will in general be a treatment involving administration of a composition that reduces the apoptotic activity of the HIP-apoptosis modulating protein. As used in the specification and claims hereof, the term "administration" includes direct administration of a composition active to reduce apoptotic activity as well as indirect administration which might include administration of pro-drugs or nucleic acids that encode the desired therapeutic composition.

One class of composition which can be used in the therapeutic methods of the invention are those compositions which interfere with the activity of HIP-apoptosis modulating proteins by binding to the proteins and mask and reduce the interaction of HIP-apoptosis modulating proteins with the death signaling complex. Within this class of



10

15

20

25

30

compositions are normal (non-expanded) huntingtin, administered, for example, via increased expression of exogenous HD genes; the HIP-binding region of huntingtin, administered via gene therapy techniques; and other DED-interacting peptides. Other DED-interacting peptides which might be used in a therapeutic method of this type include FADD (Beldin et al., *Cell* 85: 803-815 (1996)) and caspase 8 (Muzio et al., *Cell* 85: 817-827 (1996).

An alternative form of therapy involves the use of a mutant form of HIP1 or other HIP-apoptosis modulating protein from which the DED has been deleted. DED-containing proteins, including HIP1 are self-associating, and this self-association has been shown to be important for activity. (Muzio et al., *Cell* 85: 817-827 (1996). Thus, a protein with a deleted DED may compete with endogenous HIP-protein in the formation of these associations, thereby reducing the amount of apoptotically-active HIP-protein.

In addition to HIP1, we have identified a further human protein, HIP1a, from a human frontal cortex cDNA library. HIP1a is a family member of HIP1, and thus a HIP-apoptosis modulator in accordance with the invention. A partial sequence of HIP1a (the 5' portion of HIP1a remains to be characterized) is given by SEQ ID Nos. 6 and 7. The isolated and characterized portion of HIP1a shows 53% nucleotide identity and 58% amino acid conservation with HIP1 (Table 1, Figs. 2 and 3).

We have also isolated 2 mouse proteins mHIP1 and mHIP1a (SEQ. ID Nos. 8-11) which appear to be the murine homologues of human HIP1 and HIP1a. As in the case of human HIP1a, the 5' portion of mHIP1 remains to be isolated. At present, mHIP1 shows 85% nucleotide identity and 90% amino acid conservation with huHIP1 (Table 1, Figs. 2 and 3). mHIP1a shows 60% nucleotide identity and 61% amino acid conservation with huHIP1 (Table 1, Figs. 2 and 3). mHIP1a shows stronger homology to huHIP1a; it shows 87% nucleotide identity and 91% amino acid conservation with huHIP1a (Table 1, Figs. 2 and 3). Taken together these findings indicate that mHIP1 is the murine homologue of huHIP1 whereas mHIP1a is most likely the murine homologue of huHIP1a. As mentioned previously, HIP1 shows sequence similarity to Sla2p in S. cerevisiae and the hypothetical protein ZK370.3 in C. elegans. Similarly, huHIP1a, mHIP1, and mHIP1a show sequence similar to Sla2p and ZK370.3 (Table 2). The carboxy-terminal regions of huHIP1a, mHIP1, and mHIP1a all show considerable homology to the mammalian membrane

10

15

20

25

30

cytoskeletal-associated protein, talin. This suggests that these 3 proteins may also play a role in the regulation of membrane events through interactions with the underlying cytoskeleton.

HIP1 contains a death effector domain (DED), a domain which is also present in a number of proteins involved in the apoptotic pathway (Fig. 4). This suggests that HIP1 may act as a modulator of the apoptosis pathway. The DED in huHIP1 is present between amino acid positions 287 and 368. Similarly, HIP1a, mHIP1, and mHIP1a also contain a DED. In huHIP1a the DED is present at amino acids 1-78 of the recovered fragment. In mHIP1 and mHIP1a, the DED are present at amino acids 128- 210 and 388-470, respectively. The DED present in huHIP1a, mHIP1 and mHIP1a all show significant percentage amino acid conservation to the DED present in huHIP1 (Table 3).

Increasing expression of normal (non-expanded) huntingtin or the HIP-apoptotic modulator-binding portion thereof, a modified HIP-apoptotic modulator in which the DED has been deleted or of a DED-interacting protein or peptide can be accomplished using gene therapy approaches. In general, this will involve introduction of DNA encoding the appropriate protein or peptide in an expressable vector into the brain cells. Expression of HIP-apoptosis modulating proteins may also be useful in treatment of cancer in which case application to other cell types would be desired, and cells expressing HIP-apoptosis modulating proteins may be used for screening of therapeutic compounds. Thus, in a more general sense, expression vectors are defined herein as DNA sequences that are required for the transcription of cloned copies of genes and the translation of their mRNAs in an appropriate cell type. Specifically designed vectors allow the shuttling of DNA between hosts such as bacteria-yeast or bacteria-animal cells. An appropriately constructed expression vector may contain: an origin of replication for autonomous replication in host cells, selectable markers, a limited number of useful restriction enzyme sites, a potential for high copy number, and active promoters. A promoter is defined as a DNA sequence that directs RNA polymerase to bind to DNA and initiate RNA synthesis. A strong promoter is one which causes mRNAs to be initiated at high frequency. Expression vectors may include, but are not limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses.

A variety of mammalian expression vectors may be used to express recombinant



10

15

20

25

HIP-apoptosis modulating proteins or fragments thereof in mammalian cells. Commercially available mammalian expression vectors which may be suitable for recombinant HIPapoptosis modulating protein expression, include but are not limited to, pMClneo (Stratagene), pXT1 (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593) pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460), and 1ZD35 (ATCC 37565). Other vectors which have been shown to be suitable expression systems in mammalian cells include the herpes simplex viral based vectors: pHSV1 (Geller et al. Proc. Natl. Acad. Sci 87:8950-8954 (1990)); recombinant retroviral vectors: MFG (Jaffee et al. Cancer Res. 53:2221-2226 (1993)); Moloney-based retroviral vectors: LN, LNSX, LNCX, LXSN (Miller and Rosman Biotechniques 7:980-989 (1989)); vaccinia viral vector: MVA (Sutter and Moss Proc. Natl. Acad. Sci. 89:10847-10851 (1992)); recombinant adenovirus vectors: pJM17 (Ali et al Gene Therapy 1:367-384 (1994)), (Berkner K. L. Biotechniques 6:616-624 1988); second generation adenovirus vector: DE1/DE4 adenoviral vectors (Wang and Finer Nature Medicine 2:714-716 (1996)); and Adeno-associated viral vectors: AAV/Neo (Muro-Cacho et al. J. Immunotherapy 11:231-237 (1992)).

The expression vector may be introduced into host cells via any one of a number of techniques including but not limited to transformation, transfection, infection, protoplast fusion, and electroporation. The expression vector-containing cells are clonally propagated and individually analyzed to determine whether they produce the desired protein. Delivery of retroviral vectors to brain and nervous system tissue has been described in US Patents Nos. 4,866,042, 5,082,670 and 5,529,774, which are incorporated herein by references. These patents disclose the use of cerebral grafts or implants as one mechanism for introducing vectors bearing therapeutic gene sequences into the brain, as well as an approach in which the vectors are transmitted across the blood brain barrier.

To further illustrate the methods of making the materials which are the subject of this invention, and the testing which has established their utility, the following non-limiting experimental procedures are provided.

EXAMPLE 1

IDENTIFICATION OF INTERACTING PROTEINS

GAL4-HD cDNA constructs

5

10

15

20

25

An HD cDNA construct (44pGBT9), with 44 CAG repeats was generated encompassing amino acids 1 - 540 of the published HD cDNA. This cDNA fragment was fused in frame to the GAL4 DNA-binding domain (BD) of the yeast two-hybrid vector pGBT9 (Clontech). Other HD cDNA constructs, 16pGBT9, 80pGBT9 and 128pGBT9 were constructed, identical to 44pGBT9 but included only 16, 80 or 128 CAG repeats, respectively.

Another clone (DMKDBamHIpGBT9) containing the first 544 amino acids of the myotonic dystrophy gene (a gift from R. Korneluk) was fused in-frame with the GAL4-DNA BD of pGBT9 and was used as a negative control. Plasmids expressing the GAL4-BDRAD7 (D. Gietz, unpublished) and SIR3 were used as a positive control for the β -galactosidase filter assay.

The clones IT15-23Q, IT15-44Q and HAP1 were generous gifts from Dr. C. Ross. These clones represent a previously isolated huntingtin interacting protein that has a higher affinity for the expanded form of the HD protein.

Yeast strains, transformations and β-galactosidase assays

The yeast strain Y190 (MATa leu2-3,112, ura3-52, trp1-901, his3-Δ200, ade2-101, gal4Δgal80Δ, URA3::GAL-lacZ, LYS2::GAL-HIS3,cyc^r) was used for all transformations and assays. Yeast transformations were performed using a modified lithium acetate transformation protocol and grown at 30 C using appropriate synthetic complete (SC) dropout media.

The β-galactosidase chromogenic filter assays were performed by transferring the yeast colonies onto Whatman filters. The yeast cells were lysed by submerging the filters in liquid nitrogen for 15-20 seconds. Filters were allowed to dry at room temperature for at least five minutes and placed onto filter paper presoaked in Z-buffer (100 mM sodium phosphate (pH7.0) 10 mM KCl, 1 mM MgSO₄) supplemented with 50 mM



10

15

20

25

30

2-mercaptoethanol and 0.07 mg/ml 5-bromo-4-chloro-3-indolyl β -D-galactoside (X-gal). Filters were placed at 37 C for up to 8 hours.

Yeast two-hybrid screening for huntingtin interacting protein (HIP)

cDNAs from an human adult brain MatchmakerTM cDNA library (Clontech) was transformed into the yeast strain Y190 already harboring the 44pGBT9 construct. The transformants were plated onto one hundred 150 mm x 15 mm circular culture dishes containing SC media deficient in Trp, Leu and His. The herbicide 3-amino-triazole (3-AT) (25mM) was utilized to limit the number of false His+ positives (31). The yeast transformants were placed at 30 C for 5 days and β-galactosidase filter assays were performed on all colonies found after this time, as described above, to identify β-galactosidase+ clones. Primary His+/β-galactosidase+ clones were then orderly patched onto a grid on SC -Trp/-Leu/-His (25 mM 3AT) plates and assayed again for His+ growth and the ability to turn blue with a filter assay. Secondary positives were identified for further analysis. Proteins encoded by positive cDNAs were designated as HIPs (Huntingtin Interactive Proteins). Approximately 4.0 x 10⁷ Trp/Leu auxotrophic transformants were screened and of 14 clones isolated 12 represented the same cDNA (HIP1), and the other 2 cDNAs, HIP2 and HIP3 were each represented only once.

The HIP cDNA plasmids were isolated by growing the His+/β-galactosidase+ colony in SC -Leu media overnight, lysing the cells with acid-washed glass beads and electroporating the bacterial strain, KC8 (leuB auxotrophic) with the yeast lysate. The KC8 ampicillin resistant colonies were replica plated onto M9 (-Leu) plates. The plasmid DNA from M9+ colonies was transformed into DH5-a for further manipulation.

EXAMPLE 2

CONFIRMATION OF INTERACTIONS

The HIP1-GAL4-AD cDNA activated both the lac-Z and His reporter genes in the yeast strain Y190 only when co-transformed with the GAL4-BD-HD construct, but not the negative controls (Fig. 1) of the vector alone or a random fusion protein of the myotonin kinase gene. In order to assess the influence of the polyglutamine tract on the interaction

10

15

20

25

30

between HIP1 and HD, semi-quantitative β -galactosidase assays were performed. GAL4-BD-HD fusion proteins with 16, 44, 80 and 128 glutamine repeats were assayed for their strength of interaction with the GAL4-AD-HIP1 fusion protein.

Liquid β-galactosidase assays were performed by inoculating a single yeast colony into appropriate synthetic complete (SC) dropout media and grown to OD600 0.6-1.5. Five millilitres of overnight culture was pelleted and washed once with 1 ml of Z-Buffer, then resuspended in 100 ml Z-Buffer supplemented with 38 mM 2-mercaptoethanol, and 0.05% SDS. Acid washed glass beads (~100 ml) were added to each sample and vortexed for four minutes, by repeatedly alternating a 30 seconds vortex, with 30 seconds on ice. Each sample was pelleted and 10 ml of lysate was added to 500 ml of lysis buffer. The samples were incubated in a 30 C waterbath for 30 seconds and then 100 ml of a 4 mg/ml o-nitrophenyl b-D galactopyranoside (ONPG) solution was added to each tube. The reaction was allowed to continue for 20 minutes at 30 C and stopped by the addition of 500 ml of 1 M Na₂CO₃ and placing the samples on ice. Subsequently, OD420 was taken in order to calculate the β-galactosidase activity with the equation 1000 x OD420/(t x V x OD600) where t is the elapsed time (minutes) and V is the amount of lysate used.

The specificity of the HIP1-HD interaction can be observed using the chromogenic filter assay. Only yeast cells harboring HIP1 and HD activate both the HIS and lacZ reporter genes in the Y190 yeast host. The cells that contain the HIP1 with HD constructs with 80 or 128 CAG repeats turn blue approximately 45 minutes after the cells with the smaller sized repeats (16 or 44).

No difference in the β -galactosidase activity was observed between the 16 and 44 repeats or between the 80 and 128 repeats. However, a significant difference (p<0.05) in activity is seen between the smaller repeats (16 and 44) and the larger repeats (80 and 128). (Figure 1)

EXAMPLE 3

DNA SEQUENCING, cDNA ISOLATION AND 5' RACE

Oligonucleotide primers were synthesized on an ABI PCR-mate oligo-synthesizer.

DNA sequencing was performed using an ABI 373 fluorescent automated DNA sequencer.



10

15

20

25

The HIP cDNAs were confirmed to be in-frame with the GAL4-AD by sequencing across the AD-HIP1 cloning junction using an AD oligonucleotide (5'GAA GAT ACC CCA CCA AAC3'). (Seq. ID No. 12)

Subsequently, primer walking was used to determine the remaining sequences. A human frontal cortex >4.0 kb cDNA library (a gift from S. Montal) was screened to isolate the full length HIP1 gene. Fifty nanograms of a 558 base pair Eco RI fragment from the original HIP1 cDNA was radioactively labeled with [α³²P]-dCTP using nick-translation and the probe allowed to hybridized to filters containing >105 pfu/ml of the cDNA library overnight at 65°C in Church buffer (see Northern blot protocol). The filters were washed at 65°C for 10 minutes with 1 X SSPE, 15 minutes at 65°C with 1 X SSPE and 0.1% SDS, then for thirty minutes and fifteen minutes with 1 X SSPE and 0.1% SDS. The filters were exposed to X-ray film (Kodak, XAR5) overnight at -70°C. Primary positives were isolated and replated and subsequent secondary positives were hybridized and washed as for the primary screen. The resulting positive phage were converted into plasmid DNA by conventional methods (Stratagene) and the cDNA isolated and sequenced.

In order to obtain the most 5' sequence of the HIP1 gene, a Rapid Amplification of cDNA Ends (RACE) protocol was performed according to the manufacturers recommendations (BRL). First strand cDNA was synthesized using the oligo HIP1-242R (5' GCT TGA CAG TGT AGT CAT AAA GGT GGC TGC AGT CC 3'). (Seq. ID No. 13) After dCTP tailing the cDNA with terminal deoxy transferase, two rounds of 35 cycles (94°C 1 minute; 53°C 1 minute; 72°C 2 minutes) of PCR using HIP1-R2 (5' GGA CAT GTC CAG GGA GTT GAA TAC 3') (Seq. ID No. 14) and an anchor primer (5' (CUA)4 GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG3') (BRL ,Seq. ID No. 15)) were performed. The subsequent 650 base pair PCR product was cloned using the TA cloning system (Invitrogen) and sequenced using T3 and T7 primers. Sequences ID Nos. 1 and 3 show the sequence of the HIP1 cDNAs obtained.

EXAMPLE 4

DNA AND AMINO ACID ANALYSES

Overlapping DNA sequence was assembled using the program MacVector and sent via email or Netscape to the BLAST server at NIH (http://www.ncbi.nlm.nih.gov) to search for sequence similarities with known DNA (blastn) or protein (tblastn) sequences. Amino acid alignments were performed with the program Clustalw.

EXAMPLE 5

FISH DETECTION SYSTEM AND IMAGE ANALYSIS

The HIP1 cDNA isolated from the two-hybrid screen was mapped by fluorescent in situ hybridization (FISH) to normal human lymphocyte chromosomes counterstained with propidium iodide and DAPI. Biotinylated probe was detected with avidin-fluorescein isothiocyanate (FITC). Images of metaphase preparations were captured by a thermoelectrically cooled charge coupled camera (Photometrics). Separate images of DAPI banded chromosomes and FITC targeted chromosomes were obtained. Hybridization signals were acquired and merged using image analysis software and pseudo colored blue (DAPI) and yellow (FITC) as described and overlaid electronically. This study showed that HIP1 maps to a single genomic locus at 7q11.2.

20

25

30

5

10

15

EXAMPLE 6

NORTHERN BLOT ANALYSIS

RNA was isolated using the single step method of homogenization in guanidinium isothiocyante and fractionated on a 1.0% agarose gel containing 0.6 M formaldehyde. The RNA was transferred to a hybond N -membrane (Amersham) and crosslinked with ultraviolet radiation.

Hybridization of the Northern blot with b-actin as an internal control probe provided confirmation that the RNA was intact and had transferred. The 1.2 kb HIP1 cDNA was labeled using nick translation and incorporation of α³²P-dCTP. Hybridization of the original 1.2 kb HIP1 cDNA was carried out in Church buffer (0.5 M sodium phosphate buffer, pH 7.2, 2.7% sodium dodecyl sulphate, 1 mM EDTA) at 55 C overnight. Following

hybridization, Northern blots were washed once for 10 minutes in 2.0 X SSPE, 0.1% SDS at room temperature and twice for 10 minutes in 0.15 X SSPE, 0.1% SDS. Autoradiography was carried our from one to three days using Hyperfilm (Amersham) film at -70 C.

Analysis of the levels of RNA levels of HIP1 by Northern blot data revealed that the 10 kilo base HIP1 message is present in all tissue assessed. However, the levels of RNA are not uniform, with brain having highest levels of expression and peripheral tissues having less message. No apparent differences in RNA expression was noted between control samples and HD affected individuals.

10 EXAMPLE 7

5

15

20

25

TISSUE LOCALIZATION OF HIP1

Tissue localization of HIP1 was studied using a variety of techniques as described below. Subcellular distribution of HIP-1 protein in adult human and mouse brain Biochemical fractionation studies revealed the HIP1 protein was found to be a membrane-associated protein. No immunoreactivity was seen by Western blotting in cytosolic fractions, using the anti-HIP1-pep1 polyclonal antibody. HIP1 immunoreactivity was observed in all membrane fractions including nuclei (P1), mitochondria and synaptosomes (P2), microsomes and plasma membranes (P3). The P3 fraction contained the most HIP1 compared to other membrane fractions. HIP1 could be removed from membranes by high salt (0.5M NaCl) buffers indicating it is not an integral membrane protein, however, since low salt (0.1-0.25M NaCl) was only able to partially remove HIP1 from membranes, its membrane association is relatively strong. The extraction of P3 membranes with the non-ionic detergent, Triton X-100 revealed HIP1 to be a Triton X-100 insoluble protein. This characteristic is shared by many cytoskeletal and cytoskeletal-associated membrane proteins including actin, which was used as a control in this study. The biochemical characteristics of HIP1 described were found to be identical in mouse and human brain and was the same for both forms of the protein (both bands of the HIP1 doublet). HIP1 co-localized with huntingtin in the P2 and P3 membrane fractions, including the high-salt membrane extractions, as well as in the Triton X-100 insoluble residue. The subcellular distribution of HIP1 was unaffected by the

10

15

20

25

30

expression of polyglutamine-expanded huntingtin in transgenic mice and HD patient brain samples.

The localization of HIP1 protein was further investigated by immunohistochemistry in normal adult mouse brain tissue. Immunoreactivity was seen in a patchy, reticular pattern in the cytoplasm, appeared excluded from the nucleus and stained most intensely in a discontinuous pattern at the membrane. These results are consistent with the association of HIP1 with the cytoskeletal matrix and further indicate an enrichment of HIP1 at plasma membranes. Immunoreactivity occurred in all regions of the brain, including cortex, striatum, cerebellum and brainstem, but appeared most strongly in neurons and especially in cortical neurons. As described previously, huntingtin immunoreactivity was seen exclusively and uniformly in the cytosol.

The in situ hybridization studies showed HIP1 mRNA to be ubiquitously and generally expressed throughout the brain. This data is consistent with the immunohistochemical results and was identical to the distribution pattern of huntingtin mRNA in transgenic mouse brains expressing full-length human huntingtin.

Protein Preparation And Western Blotting For Expression Studies

Frozen human tissues were homogenized using a Polytron in a buffer containing 0.25M sucrose, 20mM Tris-HCl (pH 7.5), 10mM EGTA, 2mM EDTA supplemented with 10ug/ml of leupeptin, soybean trypsin inhibitor and 1mM PMSF, then centrifuged at 4,000rpm for 10' at 4 C to remove cellular debris. 100-150ug/lane of protein was separated on 8% SDS-PAGE mini-gels and then transferred to PVDF membranes. Huntingtin and HIP1 were electroblotted overnight in Towbin's transfer buffer (25 mM Tris-HCl, 0.192M glycine, pH8.3, 10% methanol) at 30V onto PVDF membranes (Immobilon-P, Millipore) as described (Towbin et al, *Proc. Nat'l Acad. Sci.(USA)* 76: 4350-4354 (1979)). Membranes were blocked for 1 hour at room temperature in 5% skim milk/ TBS (10mM Tris-HCl, 0.15M NaCl, pH7.5). Antibodies against huntingtin (pAb BKP1, 1:500), actin (mAb A-4700, Sigma, 1:500) or HIP1 (pAb HIP-pep1, 1:200) were added to blocking solution for 1 hour at room temperature. After 3 x 10 minutes washes in TBS-T (0.05% Tween-20/TBS), secondary Ab (horseradish peroxidase conjugated IgG, Biorad) was applied in blocking solution for 1 hour



at room temperature. Membranes were washed and then incubated in chemiluminescent ECL solution and visualized using Hyperfilm-ECL film (Amersham).

Generation of Antibodies

5

10

15

20

25

30

The generation of huntingtin specific antibodies GHM1 and BKP1 is described elsewhere (Kalchman, et al., *J. Biol. Chem.* 271: 19385-19394 (1996)). The HIP1 peptide (VLEKDDLMDMDASQQN, a.a. 76-91 of Seq. ID No. 2) was synthesized with Cys on the N-terminus for the coupling, and coupled to Keyhole limpet hemocyanin (KLH) (Pierce) with succinimidyl 4-(N-maleimidomethyl) cyclohexame-1-carboxylate (Pierce). Female New Zealand White rabbits were injected with HIP1 peptide-KLH and Freund's adjuvant. Antibodies against the HIP1 peptide were purified from rabbit sera using affinity column with low pH elution. Affinity column was made by incubation of HIP1 peptide with activated thio-Sepharose (Pharmacia).

Western blotting of various peripheral and brain tissues were consistent with the RNA data. The HIP1 protein levels observed was not equivalent in all tissues. The protein expression is predominant in brain tissue, with highest amounts seen in the cortex and lower levels seen in the cerebellum and caudate and putamen.

More regio-specific analysis of HIP1 expression in the brain revealed no differential expression pattern in affected individuals when compared to normal controls, with highest levels of expression seen in both controls and HD patients in the cortical regions.

EXAMPLE 8

CO-IMMUNOPRECIPITATION OF HIP1 WITH HUNTINGTIN

Confirmation of the HD-HIP1 interaction was performed using coimmunoprepitation as follows. Control human brain (frontal cortex) lysate was prepared in the same manner as for subcellular localization study. Prior to immunoprecipitation, tissue lysate was centrifuged at 5000 rpm for 2 minutes at 4 C, then the supernatant was pre-cleared by the incubated with excess amount of Protein A-Sepharose for 30 minutes at 4°C, and centrifuged at the same condition. Fifty microlitres of supernatant (500 mg protein) was incubated with or without antibodies (10 ug of anti-huntingtin GHM1 (Kalchman, et al. 1996) or anti-synaptobrevin antibody) in the total 500 ul of incubation buffer (20mM Tris-Cl

10

15

20

25

(pH7.5), 40mM NaCl, 1mM MgCl₂) for 1 hour at 4°C. Twenty microlitres of Protein A-Sepharose (1:1 suspension, for GHM1 and no antibody control) or Protein G-Sepharose (for anti-synaptobrevin antibody; Pharmacia) was added and incubated for 1 hour at 4°C. The beads were washed with washing buffer (incubation buffer containing 0.5 % Triton X-100) three times. The samples on the beads were separated using SDS-PAGE (7.5% acrylamide) and transferred to PVDF membrane (Immobilon-P, Millipore). The membrane was cut at about 150 kDa after transfer for Western blotting (as described above). The upper piece was probed with anti-huntingtin BKP1 (1/1000) and lower piece with anti-HIP1 antibody (1/300).

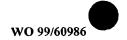
The results showed that when an anti-HIP1 polyclonal antibody was immunoreacted against a blot containing the GHM1 immunoprecipitates from the brain lysate a doublet was observed at approximately 100 kDa. When GHM1 was immunoreacted against the same immunoprecipitate the 350 kDa HD protein was also seen. The specificity of the HD-HIP1 interaction is seen as no immunoreactive bands seen are as a result of the proteins adsorbing to the Protein-A-Sepharose (Lysate + No Antibody) or when a random, non related antibody (Lysate + anti-Synaptobrevin) is used as the immunoprecipitating antibody.

EXAMPLE 9

Subcellular fractionation of brain tissue

Cortical tissue (20-100 mg/ml) was homogenized, on ice, in a 2 ml pyrex-teflon IKA-RW15 homogenizer (Tekmar Company) in a buffer containing 0.303M sucrose, 20mM Tris-HCl pH 6.9, 1mM MgCl₂, 0.5mM EDTA, 1mM PMSF, 1mM leupeptin, soybean trypsin inhibitor and 1mM benzamidine (Wood et al., *Human Molec. Genet.* 5: 481-487 (1996)).

Crude membrane vesicles were isolated by two cycles of a three-step differential centrifugation protocol in a Beckman TLA 120.2 rotor at 4 C based on the methods of Wood et al (1996). The first step precipitated cellular debris and nuclei from tissue homogenates for 5 minutes at 1300 x g (P1). The 1300 x g supernatant was subsequently centrifuged for 20 minutes at 14 000 x g to isolate synaptosomes and mitochondria (P2). Finally, microsomal



10

15

20

25

30

and plasma membrane vesicles were collected by a 35 minute centrifugation at 142 000 x g (P3). The remaining supernatant was defined as the cytosolic fraction.

High salt extraction of membranes

Aliquots of P3 membranes were twice suspended at 2mg/ ml in 0.5M NaCl, 10mM Tris-HCl, 2mM MgCl₂, pH7.2, containing protease inhibitors (see above). The same buffer without NaCl was used as a control. The membrane suspensions were incubated on ice for 30 minutes and then centrifuged at 142 000 x g for 30 minutes.

Extraction of cytoskeletal and cytoskeletal-associated proteins.

To extract cytoskeletal proteins, crude membrane vesicles from the P3 fraction membrane were suspended in a volume of Triton X-100 extraction buffer to give a protein: detergent ratio of 5:1. The composition of the Triton X-100 extraction buffer was based on the methods of Arai et al., *J. Neuroscience* 38: 348-357 (1994) and contained 2% Triton X-100, 10mM Tris-HCl, 2mM MgCl₂, 1mM leupeptin, soybean trypsin inhibitor, PMSF and benzamidine. Membrane pellets were suspended by hand with a round-bottom teflon pestle, and placed on ice for 40 minutes. Insoluble cytoskeletal matrices were precipitated for 35 minutes at 142 000 x g in a Beckman TLA 120.2 rotor. The supernatant was defined as non-cytoskeletal-associated membrane or membrane--associated protein and was removed. The remaining pellet was extracted with Triton X-100 a second time using the same conditions. We defined the final pellet as cytoskeletal and cytoskeletal-associated protein.

Solubilization of protein and analysis by SDS-PAGE and Western Blotting

Membrane and cytoskeletal protein was solubilized in a minimum volume of 1% SDS, 3M urea, 0.1mM dithiothreitol in TBS buffer and sonicated. Protein concentration was determined using the BioRad DC Protein assay and samples were diluted at least 1 X with 5 X sample buffer (250mM Tris-HCl pH 6.8, 10% SDS, 25% glycerol, 0.02% bromophenol blue and 7% 2-mercaptoethanol) and were loaded on 7.5% SDS-PAGE gels (Bio-Rad Mini-PROTEIN II Cell system) without boiling. Western blotting was performed as described above.

PCT/US99/11743

Immunohistochemistry

5

10

15

20

25

30

Brain tissue was obtained from a normal C57BL/6 adult (6 months old) male mouse sacrificed with chloroform then perfusion-fixed with 4% v/v paraformaldehyde/0.01 M phosphate buffer (4% PFA). The brain tissues were removed, immersion fixed in 4% PFA for 1 day, washed in 0.01M phosphate buffered saline, pH 7.2 (PBS) for 2 days, and then equilibrated in 25% w/v sucrose PBS for 1 week. The samples were then snap-frozen in Tissue Tek molds by isopentane cooled in liquid nitrogen. After warming to -20 C, frozen blocks derived from frontal cortex, caudate/putamen, cerebellum and brainstem were cut into 14 mm sections for immunohistochemistry. Following washing in PBS, the tissue sections were blocked using 2.5% v/v normal goat serum for 1 hour at room temperature. Primary antibodies diluted with PBS were applied to sections overnight at 4 C. Optimal dilutions for the polyclonal antibodies BKP1 and HIP1 were 1:50. Using washes of 3 x 5 minutes in PBS at room temperature, sections were sequentially incubated with biotinylated secondary antibody and then an avidin-biotin complex reagent (Vecta Stain ABC Kit, Vector) for 60 minutes each at room temperature. Color was developed using 3-3'-diaminobenzidine tetrahydrocholoride and ammonium nickel sulfate.

For controls, sections were treated as described above except that HIP1 antibody aliquots were preabsorbed with an excess of HIP1 peptide as well as a peptide unrelated to HIP1 prior to incubation with the tissue sections.

In situ hybridization

In situ hybridization was performed as previously described with some modification (Suzuki et al, *BBRC* 219: 708-713 (1996)). The RNA probes were prepared using the plasmid gt149 (Lin, B., et al., *Human Molec. Genet.* 2: 1541-1545 (1994)) or a 558 subclone of HIP1. The anti-sense and sense single-stranded RNA probes were synthesized using T3 and T7 RNA polymerases and the In Vitro Transcription Kit (Clontech) with the addition of [α³⁵S]-CTP (Amersham) to the reaction mixture. Sense RNA probes were used as negative controls. For HIP1 studies normal C57BL/6 mice were used. Huntingtin probes were tested on two different transgenic mouse strains expressing full-length huntingtin, cDNA HD10366 (44CAG) C57BL/6 mice and YAC HD10366(18CAG) FVB/N mice. Frozen brain sections



10

15

20

25

30

(10um thick) were placed onto silane-coated slides under RNase-free conditions. The hybridization solution contained 40% w/v formamide, 0.02M Tris-HCl (pH 8.0), 0.005M EDTA, 0.3 M NaCl, 0.01M sodium phosphate (pH 7.0), 1x Denhardt's solution, 10% w/v dextran sulfate (pH 7.0), 0.2% w/v sarcosyl, yeast tRNA (500mg/ml) and salmon sperm DNA (200mg/ml). The radiolabelled RNA probe was added to the hybridization solution to give 1 x 106 cpm/200 ul/ section. Sections were covered with hybridization solution and incubated on formamide paper at 65 C for 18 hours. After hybridization, the slides were washed for 30 minutes sequentially with 2x SSC, 1x SSC and high stringency wash solution (50% formamide, 2x SSC and 0.1M dithiothreitol) at 65 C, followed by treatment with RNAse A (1mg/ml) at 37 C for 30 minutes, then washed again and air-dried. The slides were first exposed on autoradiographic film (b-max, Amersham, UK) for 48 hours and developed for 4 minutes in Kodak D-19 followed by a 5 minute fixation in Fuji-fix. For longer exposures, the slides were dipped in autoradiographic emulsion (50% w/v in distilled water, NR-2, Konica, Japan), air-dried and exposed for 20 days at 4 C then developed as described. Sections were counterstained with methyl green or Giemsa solutions.

EXAMPLE 10

We determined a more precise location of the HIP1 gene on chromosome 7 in the context of a physical and genetic map of chromosome 7, and determined its genomic organization. HIP1 maps by FISH and RH mapping to chromosome band 7q11.23, which contains the chromosomal region commonly deleted in Williams-Beuren syndrome (WS). We used several methods to refine the mapping of HIP1 in this region. PCR screening of a chromosome 7-YAC-library (Scherer et al., *mammalian Genome* 3: 179-181 (1992)) with primers from the 3' UTR of HIP1 resulted in the identification of only a single positive YAC clone (HSC7E512). This YAC clone had previously been shown to map near the Williams syndrome commonly deleted region (Osborne et al., *Genomics* 45: 402-406 (1997)). The HIP1 cDNA was then used to screen a chromosome 7 specific cosmid library from the Lawrence Livermore National Laboratory (LL07NC01), and the RPCI genomic P1 derived artificial chromosome (PAC) library (Pieter de Jong, Rosswell Park, Buffalo, NY). Several PAC and cosmid clones that were already part of pre-assembled contigs in the Williams

syndrome region at 7q11.23 were identified (Fig 5). Restriction enzyme digestion, blot hybridization experiments and PCR screening confirmed that the clones contained the HIP1 gene.

We determined the exon-intron boundaries and intron sizes of HIP1. Primers were designed based on the sequence of the HIP1 transcript and used to sequence directly from the cosmid, PAC clone and long PCR products from PAC or genomic DNA. Whenever a PCR fragment generated was longer than predicted from the cDNA sequence, it was assumed to contain an intron. The size of the introns was determined by sequencing the intron directly or by PCR amplification of the introns from both genomic DNA and the cosmid or PAC clone from the region. Three sets of overlapping cosmids and a PAC clone that contain the entire coding sequence of HIP1 were characterized (Fig 5). Cosmid 181G10 and 250F2 were digested with EcoRI and cloned into the plasmid bluescript. Further sequences were generated from these plasmid subclones. Intron-exon boundary sequences were then identified by comparing HIP1 genomic and transcript sequence. The gene is contained within 75 kb and comprises 29 exons and 28 introns. The intron-exon boundary sequences are shown in Table 4, along with the exon and intron sizes. A graphic summary of these data is also shown in Fig. 5. Exons 1 to 28 contained the coding regions. The last and largest exon of the HIP1 gene was found to contain approximately 7 kb. Most of the intron-exon junctions followed the canonical GT-AG rule. An AT was found at the 3' splice site of exon 1 and an AC at the 5' splice site of exon 2. Sequence data from all the exon-intron borders of the coding region and 3'-UTR is set forth in Seq. ID Nos. 16-44. (These sequence have been deposited with GenBank as Accession Nos. AF052261 to AF052288).

Sequence analysis of previously published 5' untranslated region (GenBank accession U79734) revealed the possibility that the open reading frame extends upstream of the ATG in the exon 4 to a 5' ATG in exon 1. Although we failed to obtain any additional 5' sequences despite repeated 5' RACE analyses, an additional ATG, 284 bp upstream of the previously published exon 1 is in the same reading frame and has the surrounding sequence of TGCCATGTT which is similar to the AGCCATGGG, the consensus Kozak sequence (Kozak, M. *Nucl. Acids Res.* 15: 8125-8148 (1987)). If translated from this ATG, the protein would be highly homologous to the N-terminal portion of ZK370.3 and yeast Sla2 protein

25

5

10

15

20



10

15

20

25

30

(Fig. 6). The translated protein in the region of exons 1 to 3 shows an identity of >40% and similarity of >60% to the N-terminal part of ZK370.3. This suggests that the exons 1 to 3 are probably translated.

In western blot studies, HIP1 is identified as a 120 kd protein (11, 23), while the putative translation of the previously published cDNA gives a protein product of estimated molecular weight of approximately 100 kd. If HIP1 gene were translated from the ATG 284 bp upstream of the exon 1, the expected product would have an estimated molecular weight of 122 kd. RNA PCR studies with primers downstream of this ATG and primers in exon 7 amplify expected products of 576 and 600 bp. Taken together these data support the contention that exon 1 extends further 5' and that HIP1 gene is translated from the ATG in exon 1. Sequence analyses showed no TATA, CAAT box or any GC rich promoter sequence upstream of exon 1 ATG. The promoter prediction programs provided by the server http://dot.imgen.bcm.tmc.edu: 9331/seq.search/gene.search.html did not predict any promoter upstream of the ATG at position -284, (position 0 corresponds to the first nucleotide of published cDNA, GenBank accession U79734). This suggests that HIP1 may have additional exons.

Finally, we evaluated HIP1 gene as a candidate gene for Huntington disease in families without CAG expansion. In a large study of 1022 patients with a clinical diagnosis of HD, no CAG repeat expansion was found in 12 patients who might represent phenocopies of HD. In at least three families, linkage studies have excluded the HD locus at 4p. Mutation in an interacting protein could result in a similar phenotype as illustrated by the discovery of mutations in dystrophin associated proteins in muscular dystrophies. A mutation in HIP1 may result in altered interaction of huntingtin and HIP1 and lead to cellular toxicity as a result of more HIP1 being free in the cytosol. Thus mutations in huntingtin interacting proteins genes may cause a phenotype suggestive of HD. We studied two of the larger families diagnosed with HD without CAG expansion in HD gene, with the highly informative marker D71816 which maps centromeric and very close to HIP1 gene. The clinical findings in both the families were compatible with a diagnosis of HD, although there were atypical features. In family 1733, HIP1 locus appears to be excluded, as there are two recombinants with the marker. Individuals II-5 and II-7 who do not share the haplotype with

the affected individuals are now 41 and 39 years old and have normal neurological examinations.

In the family 1602, a lod score of 1.92 is obtained with the marker D7S1816 at θ_{max} =0. Sequencing of all the coding exons did not reveal any mutation in any exon sequence. The promoter sequence has not been examined. Subsequently a whole genome scan revealed a higher lod scores for markers on chromosome 20p.

EXAMPLE 11

A mouse brain lambda ZAPII cDNA library (Stratagene # 93609) was screened with various mouse ESTs which showed homology to the human HIP1 cDNA sequence (see Fig. 7). The ESTs were initially isolated from the non-redundant Database of GenBank EST Division by performing a BLASTN using a fragment of the human HIP1 cDNA as the query. We obtained 4 different ESTs which showed homology to HIP1: 1) aa110840 (clone 520282) which is 399bp and shows 58% identity, at the nucleotide level, to position 1880 to 2259 of the HIP1 cDNA. 2) w82687 (clone 404331) which is 420bp and shows 66% identity, at the nucleotide level, to position 2750 to 2915 of the HIP1 cDNA. 3) aa138903 (clone 586510) which is 509bp and shows 88% identity, at the nucleotide level, to position 2763 to 2832 of the HIP1 cDNA. 4) aa388714 (569088) which is 404bp and shows 88% identity, at the nucleotide level, to position 2475 to 2692 of the HIP1 cDNA.

20

25

30

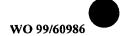
5

10

15

mHIP1:

Fifty nanograms of a 362bp KpnI & PvuII fragment of clone 569088 (containing EST aa388714) was radioactively labeled with [32-P]-dCTP using random-priming. The probe was allowed to hybridize to filters containing > 2x 10⁵ pfu/ml of the mouse brain lambda ZAPII cDNA library (Stratagene # 93609) overnight at 65°C in Church buffer (0.5M sodium phosphate buffer (pH 7.2), 2.7% SDS, 1mM EDTA). The filters were washed at room temperature for 15 minutes with 2XSSPE, 0.1% SDS, then at 65°C for 20 minutes with 1XSSPE, 0.1%SDS and finally twice at 65°C with 0.5 XSSPE, 0.1%SDS. The filters were exposed to X-ray film (Kodak, XAR5) overnight at -70 C. Primary positives were isolated, replated and subsequent secondary positives were hybridized and washed as for the primary



screen. The resulting positive phage was converted into plasmid DNA by conventional methods (Stratagene) and the cDNA termed 4n-n1, was isolated and sequenced 551bp and 541bp from the T7 and T3 end, respectively. 4n-n1 is 2.2kb in length and the T7 end showed 72% identity, at the nucleotide level, to position 1486 to 1715 of the HIP1 cDNA. The 2.2kb insert from 4n-n1 was excised using EcoR1. Fifty nanograms of the 2.2kb insert was used to produced a radioactive probe and used to screen the mouse brain lambda ZAPII cDNA library (Stratagene # 93609) in the same manner as above. The resulting positive phage was converted into plasmid DNA by conventional methods (Stratagene) and the cDNA termed mHIP1a, was isolated and completely sequenced. mHIP1 is 2.3kb in length and showed 85% identity, at the nucleotide level, to position 726 to 3072 of the HIP1 cDNA.

mHIP1a:

5

10

15

Fifty nanograms of a 1.3kb EcoRI & Ncol fragment of clone 404331 (containing EST w82687) was radioactively labeled with [32-P]-dCTP using random--priming. The probe was allowed to hybridize to filters containing > 2x 10⁵ pfu/ml of the mouse brain lambda ZAPII cDNA library (Stratagene # 93609) overnight at 65°C in Church buffer (see above). The filters were washed at room temperature for 15 minutes with 2XSSPE, 0.1% SDS, then at 65°C for 20 minutes with 1XSSPE, 0.1%SDS and finally twice at 65°C with 0.2XSSPE, 0.1%SDS. The filters were exposed to X-ray film (Kodak, XAR5) overnight at -70°C. Primary positives were isolated, replated and subsequent secondary positives were hybridized and washed as for the primary screen. The resulting positive phage was converted into plasmid DNA by conventional methods (Stratagene) and the cDNA termed mHIP1a, was isolated and completely sequenced. mHIP1a is 3.96 kb in length and shows 60% identity, at the nucleotide level, to position 12 to 2703 of the HIP1 cDNA.

25

20

EXAMPLE 12

HIP1a:

The entire mHIP1a cDNA sequence was used to screen the non-redundant Database of GenBank EST Division. We identified a human EST, T08283, which showed homology to

10

20

25

- 27 -

mHIP1a. T08383 (clone HIBBB80) is 391bp and shows 87% identity, at the nucleotide level, to position 2904 to 3113 of the mHIP1a cDNA.

Fifty nanograms of a 1.6kb HindIIII & Not1I fragment of clone 404331 (containing EST T08283) was radioactively labeled with [32-P]-dCTP using random-priming. The probe was allowed to hybridize to filters containing > 2x 105 pfu/ml of a human frontal cortex lambda cDNA library overnight at 65 C in Church buffer (see above). The filters were washed at 65 C for 10 minutes with 1XSSPE, 0.1% SDS, and then for 30 minutes and 15 minutes with 0.1XSSPE, 0.1%SDS. The filters were exposed to X-ray film (Kodak, XAR5) overnight at -70 C. Primary positives were isolated, replated and subsequent secondary positives were hybridized and washed as for the primary screen. The resulting positive phage was converted into plasmid DNA by conventional methods (Stratagene) and the cDNA termed HIP1a, was isolated and completely sequenced. HIP1a is 3.2 kb in length and shows 53% identity, at the nucleotide level, to position 876 to 3058 of the HIP1 cDNA.

15 EXAMPLE 13

Following the identification of a 1.2 kb partial human HIP-1 cDNA by yeast two-hybrid interaction studies, a 3.9 kb HIP-1 fragment was isolated from a cDNA library, ligated to a 5' RACE product then subcloned into the mammalian expression vector pCI-neo (Promega). This construct, CMV-HIP-1, expresses HIP-1 from the CMV promoter and was used in the cell expression studies described below. Mouse HIP-1a (mHIP-1a) was also subcloned into a CMV driven expression vector for cell culture expression studies.

EXAMPLE 14

Huntingtin proteins with expanded polyglutamine tracts can aggregate into large, irregularly shaped deposits in HD brains, transgenic mice and in vitro cell culture. We have shown that in HEK (human embryonic kidney) 293T cells the aggregation of full-length and larger huntingtin fragments occurs after the cells have been exposed to a period of apoptotic stress. In order to assess the consequence of HIP-1 expression in cultured cells, we used huntingtin aggregation as one marker of viability.



10

15

20

25

30

Human embryonic kidney cells (HEK 293T) were grown on glass coverslips in Dulbecco's modified Eagle medium (DMEM, Gibco, NY) with 10% fetal bovine serum and antibiotics, in 5% CO2 at 37°C. The cells were transfected at 30% confluency with the calcium phosphate protocol by mixing Qiagen-prepared DNA (Qiagen, CA) with 2.5 M CaCl₂, then incubating at room temperature for 10 min. 2X HEPES buffer (240 mM NaCl, 3.0 mM Na₂HPO₄, 100 mM HEPES, pH 7.05) was added to the DNA/calcium mixture, incubated at 37°C for 60 sec, then added to the cells. After 12-18 h, the media was removed, the cells were washed and fresh media was added. At 36 h post-transfection, the cells were exposed to an apoptotic stress by treatment with 35 uM tamoxifen (Sigma) for 1 hour, or left untreated, then processed for immunofluorescence. The cells were washed with PBS, fixed in 4% paraformaldehyde/PBS solution for 20 minutes at room temperature then permeabilized in 0.5% Triton X-100/PBS for 5 min. Following three PBS washes, the cells were incubated with anti-huntingtin antibody MAB2166 (Chemicon) (1:2500 dilution) and anti-HIP-1 antibody HIP-1fp (1:100 dilution) in 0.4% BSA/PBS for 1 h at room temperature in a humidified container. The primary antibody was removed, the cells were washed and secondary antibodies conjugated to Texas red or FITC were added at a 1:600-1:800 dilution for 30 min at room temperature. The cells were then washed again, and the coverslips were mounted onto slides with DAPI (4',6'-diamindino-2 phenylindole, Sigma) as a nuclear counter-stain. Immunofluorescence was viewed using a Zeiss (Axioscope) microscope, digitally captured with a CCD camera (Princeton Instrument Inc.) and the images were colourized and overlapped using the Eclipse (Empix Imaging Inc.) software program. Appropriate control experiments were performed to determine the specificity of the antibodies, including secondary antibody only and mock transfected cells.

The huntingtin fragment HD1955 was used in the aggregation studies. This fragment represents the N-terminal 548 amino acids of huntingtin, and corresponds approximately to the polyglutamine-containing fragment produced by caspase 3 cleavage of huntingtin.

Transfection of HD1955 with 15 polyglutamines (HD1955-15) results in a diffuse cytoplasmic distribution of the expressed protein. Transfection of HD1955 with 128 polyglutamines (HD1955-128) also results in diffuse cytoplasmic expression. However, exposure of cells transfected with HD1955-128 to tamoxifen results in a marked

10

15

20

25

redistribution of huntingtin. In 29% of cells expressing HD1955-128, the huntingtin protein appears as dense aggregates that are localized in the perinuclear area of the cell. In contrast, less than 1% of HD1955-128 expressing cells contain aggregates in the absence of tamoxifen, and 0% of HD1955-15 cells form aggregates in the presence or absence of tamoxifen treatment.

Co-transfection of HIP-1 and HD1955 was used to test the influence of HIP-1 on huntingtin aggregation. As a control, b-galactosidase was co-transfected with HD1955. In the control transfections, 1-2% of cells expressing HD1955-128 formed aggregates in the absence of tamoxifen, similar to HD1955-128 expressed alone. However, when HD1955-128 was co-expressed with HIP-1, an average of 14% of huntingtin-expressing cells contained aggregates with no tamoxifen treatment. Double-labeling demonstrated that the majority of the cells containing aggregates also expressed HIP-1, directly implicating HIP-1 in the increase in aggregation. Therefore, these results indicate that HIP-1 provides sufficient stress on the huntingtin-expressing cells to form aggregates, to the extent that tamoxifen is no longer necessary.

EXAMPLE 15

We next designed a series of experiments to identify a region of HIP-1 sufficient for inducing aggregate formation of HD1955-128. As described above, HIP-1 contains a domain with high homology to the death effector domains (DED) present in many apoptosis-related proteins. The DED domain of HIP-1 was ligated in-frame to HD1955-128, 3' from the caspase-3 cleavage site. Transfection of the resulting fusion protein with the DED ligated in the sense orientation (HD1955-128-DEDsense) resulted in a large number (30-50%) of cells containing aggregates, without tamoxifen incubation. In contrast, expression of a huntingtin-DED fusion protein with DED in the antisense orientation (HD1955-128-DEDantisense) did not have more aggregates than the HD1955-128 no tamoxifen control. Therefore, the DED domain of HIP-1 is sufficient to stress the cells, causing aggregate formation.



EXAMPLE 16

To directly assess the effect of HIP-1 expression on cell viability, mitochondrial function tests were performed on 293T cells transfected with HIP-1. The assessment of mitochondrial function, using the MTT assay (Carmichael et al., *Cancer Res.* 47: 936-942 (1987); Vistica et al., *Cancer Res.* 51: 2515-2520 (1991)), is a standard method to measure cell viability. The MTT assay quantitates the formation of a coloured substrate (formazan), with the mitochondria of viable cells forming more substrate than non-viable cells. Since decreased mitochondrial activity is an early consequence of many cellular toxins, the MTT assay provides an early indicator of cell damage.

10

15

5

For cell viability assays, HEK 293T cells were seeded at a density of 5 x 10⁴ cells into 96-well plates and transfected with 0.1 ug or 0.08 ug HIP-1 or 0.1 ug of the control construct lacZ using the calcium phosphate method described above. At 24-36 hours post-transfection tamoxifen-treated cells were incubated for 2 hours in a 1:10 dilution of WST-1 reagent (Boehringer Mannheim) and release of formazan from mitochondria was quantified at 450 nm using an ELISA plate reader (Dynatech Laboratories) (Carmichael et al., 1987; Mosmann, *J. Immunol. Meth* 65: 55-63 (1983)). One way ANOVA and Newman-Keuls test were used for statistical analysis. The transfection efficiency, measured by β-galactosidase staining and immunofluoresence, was approximately 50%.

20

25

30

We have previously demonstrated that expression of mutant huntingtin results in increased susceptibility to an apoptotic stress induced by sub-lethal doses of tamoxifen in transfected 293T cells (Martindale et al., 1998). A similar assay was used to test the consequence of HIP-1 expression. With 0.1 ug transfected HIP-1 DNA, after 24 hr expression, HIP-1 resulted in increased cell death in response to tamoxifen, compared with the tamoxifen-treated β-galactosidase control (p<0.01, n=4). Reducing the amount of transfected HIP-1 DNA to 0.08 ug also resulted in increased cell death compared with control (p<0.01, n=4), indicating the high potency of HIP-1 (Fig. 8). Furthermore, increased cell death in cells transfected with HIP-1 was observed in the absence of apoptotic stress at 48 hrs post-transfection, but was so severe that is could not be accurately quantitated. Thus, an earlier time point (24 hr) had to be used for better reproducibility, using an apoptotic stress to unmask the phenotype.

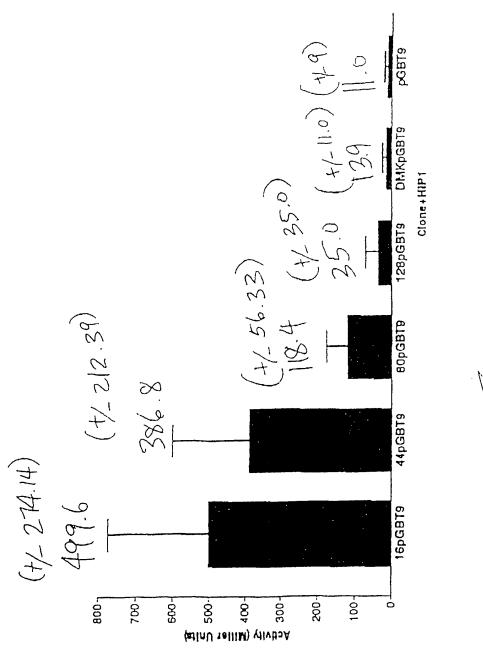
In order to model a pathogenic interaction of HIP-1 and huntingtin in the HEK 293 mammalian cell system, HIP-1 was transfected into cell lines stably expressing huntingtin. Two cell lines were chosen for the initial studies, one line expressed the truncated HD1955 construct with 15 glutamines, and the second expressed the HD1955 with 128 repeats. Western blotting indicated that the cell lines expressed huntingtin at similar levels. To assess whether HIP-1 is toxic in the presence of mutant huntingtin, 0.1 ug HIP-1 DNA was transfected into the HD1955-128 cell line. Transfection of HIP-1 into the HD1955-15 cell line was used as the wild-type huntingtin control, and transfection of LacZ into both cell lines was the control for transfection-related toxicity (Figs 9A and 9B). MTT toxicity assays showed that HIP-1 in the presence of mutant huntingtin (HD1955-128) was significantly more toxic than HIP-1 with wild-type huntingtin (HD1955-15), p<0.001, n=4 (Fig. 9C). This toxicity was observed at 24 hr and 36 hr post-transfection. No tamoxifen was needed to unmask the phenotype, suggesting that the combined cell stress of HIP-1 with truncated huntingtin was sufficient to reduce cell viability over control.



CLAIMS

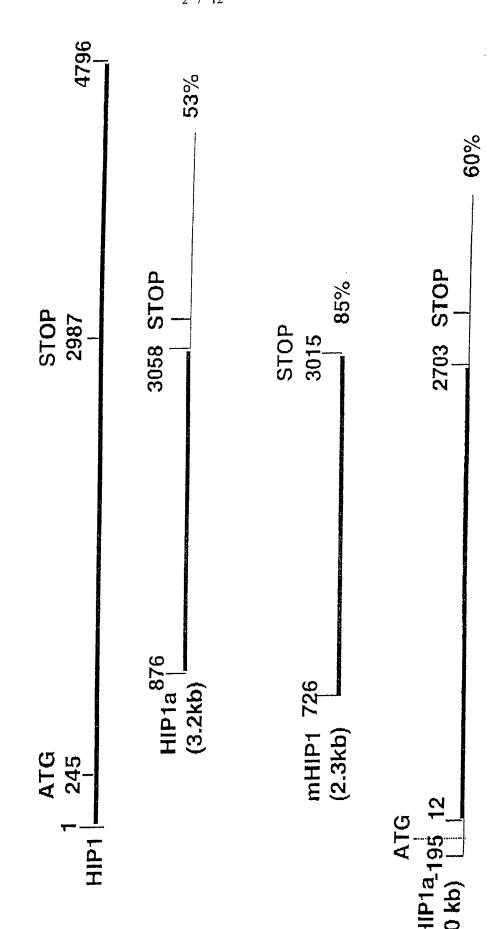
1	1.	A polypeptide comprising the sequence given by Seq. ID. No. 5.
1	2.	A cDNA molecule comprising the sequence given by Seq. ID No. 6.
1	3.	A polypeptide comprising the sequence given by Seq. ID No. 7.
1	4.	A method for ameliorating the effects of Huntington's disease in a
2	patient expressing a	HIP-apoptosis modulating protein, comprising the step of administering
3	the patient a theraped	atic composition which reduces the activity of the HIP-apoptosis
4	modulating protein.	
1	5.	A method according to claim 4, wherein the composition comprises a
2	material which binds	to HIP-apoptosis modulating protein.
1	6.	The method according to claim 4, wherein the composition comprises
2	an expression vector	encoding huntingtin having a normal number of repeats.
1	7.	An expression vector for expression of a gene in a mammalian host
2	comprising a region	encoding an HD-interacting polypeptide.
1	8.	The expression vector according to claim 7, wherein the HD-
2	interacting polypepti	de is an HIP-apoptosis modulating protein.
1	9.	The expression vector according to claim 8, wherein the HIP-apoptosis
2	modulating protein h	as a sequence which includes the amino acid sequences given by SEQ
3	ID Nos. 2, 4, 5 or 7.	

1	10. The	e expression vector of claim 7, wherein the HD-interacting				
2	polypeptide interacts differently with expanded Huntingtin than with Huntingtin having a					
3	CAG repeat region containing 15 to 35 repeats.					
1	11. The	e expression vector according to claims of claims 7-10, further				
2	comprising a region encoding Huntingtin having a polyglutamine tract of 35 or fewer.					
1	12. A r	nethod for inducing apoptotic death in cells, comprising the step of				
2	introducing into the cells an expression vector encoding at least the death effector domain of					
3	a HIP-apoptosis modulating protein whereby the death effector domain is expressed by the					
4	cells.					
1	13. The	e method of claim 12, wherein the expression vector encodes the				
2	amino acid sequence given by Seq. ID. No. 2.					
1	14. Th	e method of claim 12, wherein the expression vector encodes the				
2	amino acid sequence given by Seq. ID. No. 4.					
1	15. A 1	method for screening a composition for the ability to inhibit				
2	apoptosis induced by an HIP-apoptosis modulating protein, comprising simultaneously					
3	exposing a population of cells to the composition and an HIP-apoptosis modulating protein					
4	and measuring the extent of cell death.					

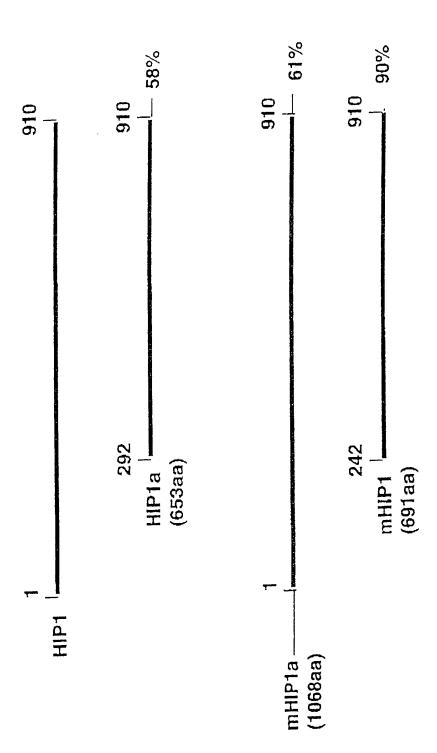


F19. 1

Fy 2 HIP1 Clones: Nucleotide Alignment



جُنِع HIP1 Clones: Protein Alignment



Fy4

>Usurpin A

SAEVIHQVEEALDTDEKKMLLFLCRDVAIDVVPPNVRDLLDILRERGKLSVCDLAELLYRVHRFDLLKRILK

>Usurpin, B

YRVLMAHIGEDLDKSDVSSLIFLMKDYMGRGKISKHKSFLDLVVELHKLNLVAPDQLDLLEKCLKNIHRIDLKTKIQK

>Casp-8 A

FSRNLYDIGELQDSEDLASLKELSLDYIPQRKOEPIKDALMIFQRLOEKRMLEESNLSFLKELLFRINRLDLLITYLN

>Casp-8 B

YRVMLYQISEEVSREELRSFKFLLQHEISKCKLDDDMNLLDIFIEMEKRVILGEGKLDILKRVCAQINKSLLKIND

>Casp-10 A

FRHKLLTIDSNLGVQDVENLKFLCIGLVPNKKLEKSSSASDVFEHLLAHDLLSEEDPFFLAELLYIIRQKKLLQHLNC

>Casp-10 B

FRNLLYELSEGIDSENLKDMIFLLKDSLPKTEMTSLSFLAFLEKQGKIDEDNLTCLEDLCKTVVPKLLRNIEK

>FADD

FLVLLHSVSSSLSSSELTELKFLCLGRVGKRKLERVQSGLDLFSMLLEQNDLEPGHTELLRELLASLRRHDLLRRVDD

>MC159 A

SLPFLRHLLEELDSHEDSLLLFLGHDAAPGCTTVTQALCSLSQQRKLTLAALVEMLYVLQRMDLLKSRFG

>MC159 B

YHKLMYCYGEELDSSELRALRLFACNLNPSLSTALSESSRPVELVLALENYGLVSPSSYSYLADMLRTLRRLDLCQQLVE

>E8

FRCLMALVNDFLSDKEVEEHYFLCAPRLESHLEPGSKKSFLRLASLLEDLELLGGDKLTFLRHLLTTIGRADLVKNLQV

>KS orfk13A

TYEVLCEVARKLGTDDREVVLFLLNVFLPQPTLAQLIGALRALKEEGRLTFPLLAECLPRAGRRDLLRDLLH

>KS orfkl3B

YQLTVLHVDGELCARDIRSLIFLSKDTIGSRSTPQTFLHNVYCMENLDLLGPTDVDALMSMLRSLSRVDLQRQVQT

>HIP1

SELEADLAEQQHLRQQAADDCEFLRAELDELRRQREDTEKAQRSLSEIERKAQANEQRYSKLKEKYSELVQNHADLLRKN

>HIP1a

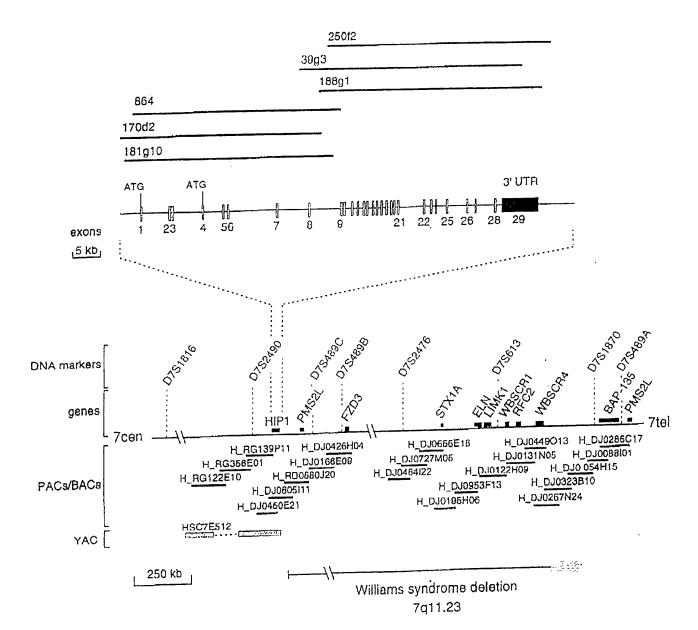
GELEEQRKQKQKALVDNEQLRHELAQLRAAQLERERSQGLREEAERKASATEARYNKLKEKHSELVHVHAELLRKNAD

>mHIPla

NGLEAELEEQRKQKQKALVDNEQLRHELAQLKALQLEGARNQGLREEAERKASATEARYSKLKEKHSELINTHAELLRKN AD

>mHIP1

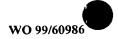
SELEAELA'EQQHLGRQAMDDCEFLRTELDELKRQREDTEKAQRSLTEIERKAQANEQRYSKLKEKYSELVQNHADLLRKN AE





\$ 1 d

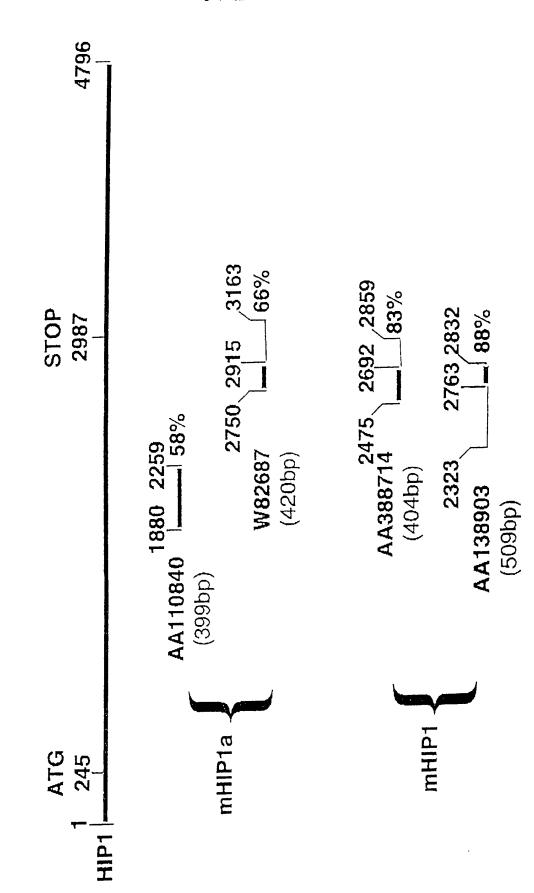
90 12	180	269	355 277	ት ድ ዓ ት ፍ	O 60 10 10 10 10	617	702 569	792	882
SQRLTHLTLLSDPLL HDHRAQAREVFY	RY H S DSLRYRNELSDMSRM ETYRYVNRPTQLSQF	V. Vensldhsrsvsvta Vyendhslrhnslip	IL LP PNFL S IPQLPENPPNFLRAS IPTLPSHAPNFLQQS	L K E ERLYRBISGLKAQLE VMLSRAVBDEKFAKE	B R K K Y BQ-Rysklkbkyse- EE-Rfnkmkguyek.	QVLQGSLETSAQSEA EGBAQAVDZKRTQLV	P NQLSEPPLIS SITYPP	CK A CKQYGRETLAYIASL CKKVLAAAKVAF	IR RA G LEVNE IL RIBBALSKSRAGDIG VKLEVNERILGCCTS
	. HKLLRDGEP V VFHKLLRDGEPNVLK I LVHKLLRDGHRKVPK 1	M L V. R S. EMFDYLECELNIEGT VENSIDHSRSVSVTA DYLLOGYDALLVIQDR VYENGNISLRANSLIP	SSNLQTFK L. FYRSSNLQTFKRLIQ YEESSNLQTFKYLVS		A . E ERKA A AQRSLSEIBRKAQAN ARVKRAELKATAA	S LESLKORLATSCREL QVLQGSLETSFQSEA BAGRALTKA EGDAGAVDZHRTQLV	E Q A VGSRKAAEQVIQDAL TQMFDHCEDALQNAT	a Clrappepadslyba Aytasiesyegundo	
QPAGSWERCPPLPPA GRLQGIDHPWGWGRL AGGGERGSLWEGLSH	L VL WEPCH LPLSSNAVLCWKFCH 1 IQLRKHPVLTWKFCH 1	E D. N F T EAGESDVANFRQLIV TL-EGDLDMAFENTI	ER RF ? K ERDKFMRQ?TELKDL HRSRFRTI?ERTKKF	DDIFGSSPSSDFNF NSQNGMNDEEDIL LNLAEAEPQYASP SSQPDEREQI	B R T AZLDELRRQRBDTBK NELALRDASRTQTDD	QR ISDZGQRKTQEÇLEV KVBEAQR	H B Q A T. IVSGAAHRBBBLSAL RKELQDTGLKLASTB ESMCQLAKDAKDQRKOLL VGSRKAABQVIQDAL NQLEEPPLIS HURBSHAMQLVQSSAEBTNKIR LABLEVAKES-GVGI TQMFDHCEDALQNAT SITYPP	EL S A LAELTSDAIMGATT AGHLLSTTLSBAASA	L D EEGSLENADSTAME NCLSKIKAIGEELLP RGLDIKQEELGDL/D KEMARITSAAIETATA
PAGSWERCPPLPPA (E KK FW V BYEKGAQTF#SVYVR EKEKSSGI?#HTYGR	P FG L D QL PRPPGNLOMSDRQLD PVVPGKLDLMDSQLK	ES . D L G ESCLPADTLQG ESQVPPDALEG	D M B MDMDASÇONLFDNKF DGTSLMGHEGEL	L QHLRQQNADOCBFLR KRZADZNRSENQRLK	XQ. DARVTXQVSHARQAQV DLERBKELEDSLER GDIQXQLEASEESK? DKDESITALNR	Q RKELQDTQLKLASTB -QLVQSSAEBTNKIR	I B L A SSCIEQLBKSWSQYL ACPRDISGLIHSITL LVNILLSNRR-LDBPL ATTONYF	RGIDIKQEELABLYD
GTANARQMCPLPQDA (K KHART I.GT VAVKEKHARTCILGT VPLXVPKHARTIIVGT	Y KLL. YEKLERTKMBYBTEN YCKLLIORVTFRYKY	S YDY OK . BKL CSHLYDYTVELLPKL TSKEYDYLVKMIFKL	S B B B A B ASSPDSEPVLEKOOL KOMDASÇQNLFDNKF GSEE	B Q P P P				ALLEKIKAIGEELIA
MLLCQGSENRRDQQL C	Q KAI E VPQTVSINKAINTQ3 V RAQUEAVQKAITYGME V	W HL GYG WGHLS-EGYGQLCSI WKHLNTSGYGPCIES	GQC PLI .ILD AGGCRLAPLIQVILD QGQCHLSPLIIAILD	L P B ALSZHISPVVVIPAE DLESYRIZHAYTHSE		H L LVQNBADLLRKN PRSBEVLALTKL	E EL D HYLARZARLEKERDS KAUIRVEELKRIID-	HL CAGSADHLLSTVTSI	
1 hip1 2 zk370.3	1 hipi 2 2k370.3	1 hip1 2 zk370.3	1 hipl 2 zk370.3	1 hip1 2 zk370.3	1 hip1 2 2k370.3	1 hip1 2 =k370.3	1 hi <u>p</u> 1 2 zk370.3	1 bipl 2 zk370.3	l bipl



<u></u>
0,
\wedge
0
6
9

958 817	1056 305	
LM .I LAS L Q ZIV SP KPY N MT3GLIS A KAVG A V AD VV GKGK F R L V BIAASTA I bipl LMQAIQVLIVASKOL QRKIVESGRGIASPK KPKARKSKMTSGLIS ASKAVGNGATVHYDA ADLVYGGKGKF KELMYCSHBIAASTA 2 zkljo.3 lmsvikglytaskel Qtrivaagkaggera Bpkkrhhymtsglis aakavgvarruss adgvytgkgkF kflivaaqgelaasta	OL . S VKADKDSS L L A VMQ TA VVA G . D?S L K KYV I BIR L KR KL 1 hipl quvaaskvkadkdsp nlaqlqqasrgvmqa tagvvastisgksqi betdmimpssmtl tqikrqeedsqvrvl blenbilgerqilge 2 zhijo.i qlevsspvkadkds Kldnisvankavmqn tagvvaavkqqqttl ndbgsldpsylsl haakkeedesqvqqu bleqselnqeraklaa	
kpy in mysglis Rpyaknskmysglis Rpykrhymysglis	TA VVA TAGVVASTISGKSQI TAGVVAAVKOQCTTL	TEKE 1090
Q ZIV SP QPRIVESGRGIASPK QTRIVAAGKAGGEPA	L L A VNQ NLAZLQQASRGVNQA KLDALSVARKAVNQN	ebgizaspptloevv VS?
EM .I EAS L Q EIV EMQAIQVELYASKOE QREIV LHSVIKQLVIASREE QTRIV	QL . S VEADEDSS L L A VNQ TA VVA G . QLVAASKVEADEDSP NLAQLQQASRGVNQA TAGVVASTISGKSQI QLEVSSRVKADEOSS KLDNESVARKAVNGN TAGVVAAVKOGTTTL	LRK BY A . 1 hipl lrkkhyelagyabgw eegisaspptloevy teke 2 zk37c.3 lrkqhymbalvamk vsp
1 bipl 2 zk370.3	1 hip1 2 zk370.3	1 hipi 2 skipc.s

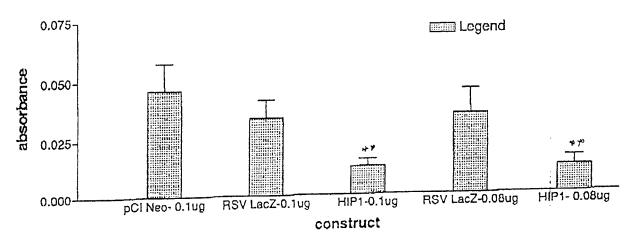
Fig 7 Mouse ESTs



mthlp1.pzm:Graph-2 - Tuo Apr 28 11:30:41 1993

Hip I increase the susceptibility to cell Stress.

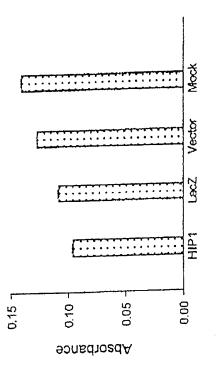
HIP1 TOXICITY



Fy8

HIP1 transfected into HD1955-15 stable cell line 36 hr post-tansfection

Hy-1 is toxic in the presence of



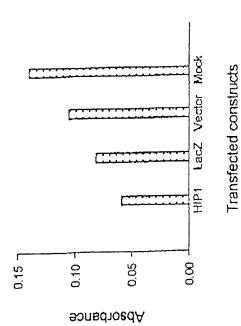
Transfected constructs

Fy 91

HIP1 transfected into HD1955-128 stable cell line

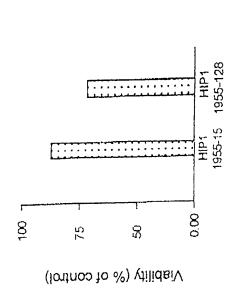
Hig-1 is toxic in the presence of hundre

36 hr post-tansfection



Fy 98

Polyglutamine-dependence of HIP-1 toxicity



Transfected constructs/cell lines



SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: Kalchman, Michael

Hayden. Michael R.

Hackam, Abigail

Chopra, Vikramjit Singh

Nicholson, Donald W.

Vallaincourt, John P.

Rasper, Dita M.

(ii) TITLE OF INVENTION: Apoptosis Modulators That Interact with the

Huntington's Disease Gene

- (iii) NUMBER OF SEQUENCES: 44
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Oppedahl & Larson
- (B) STREET: PO Box 5270
- (C) CITY: Frisco
- (D) STATE: CO
- (E) COUNTRY: USA
- (F) ZIP: 80443-5270
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: MS DOS 5.0
- (D) SOFTWARE: WordPerfect
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Larson, Marina T.
- (B) REGISTRATION NUMBER: 32038
- (C) REFERENCE/DOCKET NUMBER: UBC.P-013US2
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (970) 668-2050
- (B) TELEFAX: (970) 668-2052
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1164
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no

- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: cDNA for Huntingtin-interacting protein
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACAGCTGACA	CCCTGCAAGG	CCACCGGGAC	CGCTTCATGG	AGCAGTTTAC	50
AAAGTTGAAA	GATCTGTTCT	ACCGCTCCAG	CAACCTGCAG	TACTTCAAGC	100
GGGTCATTCA	GATCCCCCAG	CTGCCTGAGA	ACCCACCCAA	CTTCCTGCGA	150
GCCTCAGCCC	TGTCAGAACA	TATCAGCCCT	GTGGTGGTGA	TCCCTGCAGA	200
GGCCTCATCC	CCCGACAGCG	AGCCAGTCCT	AGAGAAGGAT	GACCTCATGG	250
ACATGGATGC	CTCTCAGCAG	AATTTATTTG	ACAACAAGTT	TGATGACNTC	300
TTTGGCAGTT	CATCCAGCAG	TGATCCCTTC	AATTTCAACA	GTCAAAATGG	350
TGTGAACAAG	GATGAGAAGG	ACCACTTAAT	TGAGCGACTA	TACAGAGAGA	400
TCAGTGGATT	GAAGGCACAG	CTAGAAAACA	TGAAGACTGA	GAGCCAGCGG	450
GTTGTGCTGC	AGCTGAAGGG	CCACGTCAGC	GAGCTGGAAG	CAGATCTGGC	500
CGAGCAGCAG	CACCTGCGGC	AGCAGGCGGC	CGACGACTGT	GAATTCCTGC	550
GGGCAGAACT	GGACGAGCTC	AGGNGGCAGC	GGGAGGACAC	CGAGAAGGCT	600
CAGCGGAGCC	TGTCTGAGAT	AGAAAGGAAA	GCTCAAGCCA	ATGAACAGCG	650
ATATAGCAAG	CTAAAGGAGA	AGTACAGCGA	GCTGGTTCAG	AACCACGCTG	700
ACCTGCTGCG	GAAGAATGCA	GAGGTGACCA	AACAGGTGTC	CATGGCCAGA	750
CAAGCCCAGG	TAGATTTGGA	ACGAGAGAAA	AAAGAGCTGG	AGGATTCGTT	800
GGAGCGCATC	AGTGACCAGG	GCCAGCGGAA	GACTCAAGAA	CAGCTGGAAG	850
TTCTAGAGAG	CTTGAAGCAG	GAACTTGGCA	CAAGCCAACG	GGAGCTTCAG	900
GTTCTGCAAG	GCAGCCTGGA	AACTTCTGCC	CAGTCAGAAG	CAAACTGGGC	950
AGCCGAGTTC	GCCGAGCTAG	AGAAGGAGCG	GGACAGCCTG	GTGAGTGGCG	1000
CAGCTCATAG	GGAGGAGGAA	TTATCTGCTC	TTCGGAAAGA	ACTGCAGGAC	1050
ACTCAGCTCA	AACTGGCCAG	CACAGAGGAA	TCTATGTGCC	AGCTTGCCAA	1100
AGACCAACGA	AAAATGCTTC	TGGTGGGGTC	CAGGAAGGCT	GCGGAGCAGG	1150
TGATACAAGA	CGCG				1164

- (2) INFORMATION FOR SEQ ID NO:2:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 386
- (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: Huntingtin-interacting protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Thr Ala Asp Thr Leu Gln Gly His Arg Asp Arg Phe Met Glu Gln 1 5 10 15

Phe Thr Lys Leu Lys Asp Leu Phe Tyr Arg Ser Ser Asn Leu Gln 20 25 30

Tyr Phe Lys Arg Val Ile Gln Ile Pro Gln Leu Pro Glu Asn Pro

()	
WO 99/60986		

	\mathbf{w} \mathbf{u}	00280	_										101.	
				35					40					45
Pro	Asn	Phe	Leu	Arg 50	Ala	Ser	Ala	Leu	Ser 55	Glu	His	Ile	Ser	Pro 60
Val	Val	Val	Ile	Pro 65	Ala	Glu	Ala	Ser	Ser 70	Pro	Asp	Ser	Glu	Pro 75
Val	Leu	Glu	Lys	Asp 80	Asp	Leu	Met	Asp	Met 85	Asp	Ala	Ser	Gln	Gln 90
Asn	Leu	Phe	Asp	Asn 95	Lys	Phe	Asp	Asp	Phe 100	Gly	Ser	Ser	Ser	Ser 105
Ser	Asp	Pro	Phe	Asn 110	Phe	Asn	Ser	Gln	Asn 115	Gly	Val	Asn	Lys	Asp 120
Glu	Lys	Asp	His	Leu 125	Ile	Glu	Arg	Leu	Tyr 130	Arg	Glu	Ile	Ser	Gly 135
Leu	Lys	Ala	Gln	Leu 140	Glu	Asn	Met	Lys	Thr 145	Glu	Ser	Gln	Arg	Val 150
Val	Leu	Gln	Leu	Lys 155	Gly	His	Val	Ser	Glu 160	Leu	Glu	Ala	Asp	Leu 165
Ala	Glu	Gln	Gln	His 170	Leu	Arg	Gln	Gln	Ala 175	Ala	Asp	Asp	Cys	Glu 180
Phe	Leu	Arg	Ala	Glu 185	Leu	Asp	Glu	Leu	Arg 190	Gln	Arg	Glu	Asp	Thr 195
Glu	Lys	Ala	Gln	Arg 200	Ser	Leu	Ser	Glu	Ile 205	Glu	Arg	Lys	Ala	Gln 210
Ala	Asn	Glu	Gln	Arg 215	Tyr	Ser	Lys	Leu	Lys 220	Glu	Lys	Tyr	Ser	Glu 225
Leu	Val	Gln	Asn	His 230	Ala	Asp	Leu	Leu	Arg 235	Lys	Asn	Ala	Glu	Val 240
Thr	Lys	Gln	Val	Ser 245	Met	Ala	Arg	Gln	Ala 250	Gln	Val	Asp	Leu	Glu 255
Arg	Glu	Lys	Lys	Glu 260		Glu	Asp	Ser	Leu 265	Glu	Arg	Ile	Ser	Asp 270
Gln	Gly	Gln	Arg	Lys 275		Gln	Glu	Gln	Leu 280	Glu	Val	Leu	Glu	Ser 285
Leu	Lys	Gln	Glu	Leu	Gly	Thr	Ser	Gln	Arg	Glu	Leu	Gln	Val	Leu

WO 99/60986				PCT/US99/11743
-------------	--	--	--	----------------

	wo	99/	60986											1 C1/	03///11/45
					290					295					300
Gl	n Gl	У	Ser	Leu	Glu 305	Thr	Ser	Ala	Gln	Ser 310	Glu	Ala	Asn	Trp	Ala 315
Al	a Gl	.u	Phe	Ala	Glu 320	Leu	Glu	Lys	Glu	Arg 325	Asp	Ser	Leu	Val	Ser 330
G1	y Al	.a	Ala	His	Arg 335	Glu	Glu	Glu	Leu	Ser 340	Ala	Leu	Arg	Lys	Glu 345
Le	u Gl	.n	Asp	Thr	Gln 350	Leu	Lys	Leu	Ala	Ser 355	Thr	Glu	Glu	Ser	Met 360
Су	s Gl	n	Leu	Ala	Lys 365	Asp	Gln	Arg	Lys	Met 370	Leu	Leu	Val	Gly	Ser 375
Ar	g Ly	/S	Ala	Ala	Glu 380	Gln	Val	Ile	Gln	Asp 385	Ala 386				

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4796
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: cDNA for Huntingtin-interacting protein
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CAGTGTACGG	TTGATCATAT	AACGCCGCGG	GCGGGGATTG	GTTTATATAT	50
CGCAAATTGA	TNTAGGGGGG	GGGGGATGGN	CAGAGATTTC	GCTTCATTAG	100
GCCATTATAA	GCAGGAAGGG	TTTCAAGGAA	AAAAACCCAG	AAAGTGCATA	150
TTGCACCCAC	CATGAGAAAG	GGGCAACAGA	CCTTNTGTTN	TGTTNTCAAC	200
CGCCTGCTTC	${\tt TGTTTTAGCA}$	ACGCAGTGTT	TTGGTGGAAG	TTGTGCCATG	250
TGTTCCACAA	ANTCTTCCGA	GATGGACACC	CGAACGTCCT	GAAGGACTTT	300
GTGAGATACA	GAAATGAATT	GAGTGACATG	AGCAGGATGT	GGGGCCACCT	350
GAGCGAGGGG	TATGGCCAGC	TGTGCAGCAT	CTACCTGAAA	CTGCTAAGAA	400
CCAAGATGGA	GTACCACACC	AAAAATCCCA	GGTTCCCAGG	CAACCTGCAG	450
ATGAGTGACC	GCCAGCTGGA	CGAGGCTGGA	GAAAGTGACG	TGAACAACTT	500
TTTCCAGTTA	ACAGTGGAGA	TGTTTGACTA	CCTGGAGTGT	GAACTCAACC	550
TCTTCCAAAC	AGTATTCAAC	TCCCTGGACA	TGTCCCGCTC	TGTGTCCGTG	600
ACGGCAGCAG	GGCAGTGCCG	CCTCGCCCCG	CTGATCCAGG	TCATCTTGGA	650
CTGCAGCCAC	CTTTATGACT	ACACTGTCAA	GCTTCTCTTC	AAACTCCACT	700
CCTGCCTCCC	AGCTGACACC	CTGCAAGGCC	ACCGGGACCG	CTTCATGGAG	750

WO 99/60986		PCT/US99/11743	

		TCTGTTCTAC			800
		TCCCCCAGCT			850
TCCTGCGAGC			TCAGCCCTGT	GGTGGTGATC	900
CCTGCAGAGG		CGACAGCGAG	CCAGTCCTAG	AGAAGGATGA	950
CCTCATGGAC		CTCAGCAGAA	· · · -	AACAAGTTTG	1000
ATGACATCTT			ATCCCTTCAA	TTTCAACAGT	1050
CAAAATGGTG		TGAGAAGGAC	CACTTAATTG	AGCGACTATA	1100
CAGAGAGATC		AGGCACAGCT	AGAAAACATG	AAGACTGAGA	1150
GCCAGCGGGT		CTGAAGGGCC		GCTGGAAGCA	1200
		CCTGCGGCAG	CAGGCGGCCG	ACGACTGTGA	1250
ATTCCTGCGG	GCAGAACTGG	ACGAGCTCAG	GAGGCAGCGG	GAGGACACCG	1300
	GCGGAGCCTG	TCTGAGATAG	AAAGGAAAGC	TCAAGCCAAT	1350
GAACAGCGAT	ATAGCAAGCT	AAAGGAGAAG	TACAGCGAGC	TGGTTCAGAA	1400
CCACGCTGAC	CTGCTGCGGA	AGAATGCAGA	GGTGACCAAA	CAGGTGTCCA	1450
TGGCCAGACA	AGCCCAGGTA	GATTTGGAAC	GAGAGAAAAA	AGAGCTGGAG	1500
GATTCGTTGG	AGCGCATCAG	TGACCAGGGC	CAGCGGAAGA	CTCAAGAACA	1550
GCTGGAAGTT	CTAGAGAGCT	TGAAGCAGGA	ACTTGGCACA	AGCCAACGGG	1600
AGCTTCAGGT	TCTGCAAGGC	AGCCTGGAAA	CTTCTGCCCA	GTCAGAAGCA	1650
AACTGGGCAG	CCGAGTTCGC	CGAGCTAGAG	AAGGAGCGGG	ACAGCCTGGT	1700
GAGTGGCGCA	GCTCATAGGG	AGGAGGAATT	ATCTGCTCTT	CGGAAAGAAC	1750
TGCAGGACAC	TCAGCTCAAA	CTGGCCAGCA	CAGAGGAATC	TATGTGCCAG	1800
CTTGCCAAAG	ACCAACGAAA	AATGCTTCTG	GTGGGGTCCA	GGAAGGCTGC	1850
GGAGCAGGTG	ATACAAGACG	CCCTGAACCA	GCTTGAAGAA	CCTCCTCTCA	1900
TCAGCTGCGC	TGGGTCTGCA	GATCACCTCC	TCTCCACGGT	CACATCCATT	1950
TCCAGCTGCA	TCGAGCAACT	GGAGAAAAGC	TGGAGCCAGT	ATCTGGCCTG	2000
CCCAGAAGAC	ATCAGTGGAC	TTCTCCATTC	CATAACCCTG	CTGGCCCACT	2050
TGACCAGCGA	CGCCATTGCT	CATGGTGCCA	CCACCTGCCT		2100
CCTGAGCCTG		GACCGAGGCC		ATGGCAGGGA	2150
AACCCTCGCC	TACCTGGCCT	CCCTGGAGGA	AGAGGGAAGC	CTTGAGAATG	2200
CCGACAGCAC	AGCCATGAGG	AACTGCCTGA	GCAAGATCAA	GGCCATCGGC	2250
GAGGAGCTCC	TGCCCAGGGG	ACTGGACATC	AAGCAGGAGG	AGCTGGGGGA	2300
CCTGGTGGAC	AAGGAGATGG	CGGCCACTTC	AGCTGCTATT	GAAACTTGCA	2350
CGGCCAGAAT	AGAGGAGATG	CTCAGCAAAT	CCCGAGCAGG	AGACACAGGA	2400
GTCAAATTGG	AGGTGAATGA	AAGGATCCTT			2450
		TCGTGGCCTC			2500
		ACAGCATCCC			2550
		ACTTATCTCA			2600
		ATGCAGCTGA			2650
		GTGTGTTCTC			2700
		CAAGGTGAAA			2750
		CCTCTCGGGG			2800
		TCCGGCAAAT			2850
		GACGCTGACA			2900
		TAGAGCTAGA			2950
		CGGAAAAAGC		_	3000
		AACAGAGGCA			
		AGAGCCAAAC			3050
		CGTGTGTGTT			3100
		CAGCCACACC			3150
					3200
		CTGTTCTTTT			3250
U111GUACCC	111001CATC1	CIGIICIIII		TAGTIAGCAT	3300

CCAGGCTGGC	CAGTGCTGCC	CATGAGCAAG	CCTAGGTACG	AAGAGGGGTG	3350
GTGGGGGGCA	GGGCCACTCA	ACAGAGAGGA	CCAACATCCA	GTCCTGCTGA	3400
CTATTTGACC	CCCACAACAA	TGGGTATCCT	TAATAGAGGA	GCTGCTTGTT	3450
GTTTGTTGAC	AGCTTGGAAA	GGGAAGATCT	TATGCCTTTT	CTTTTCTGTT	3500
TTCTTCTCAG	TCTTTTCAGT	TTCATCATTT	GCACAAACTT	GTGAGCATCA	3550
GAGGGCTGAT	GGATTCCAAA	CCAGGACACT	ACCCTGAGAT	CTGCACAGTC	3600
AGAAGGACGG	CAGGAGTGTC	CTGGCTGTGA	ATGCCAAAGC	CATTCTCCCC	3650
CTCTTTGGGC	AGTGCCATGG	ATTTCCACTG	CTTCTTATGG	TGGTTGGTTG	3700
GGTTTTTTGG	TTTTGTTTTT	TTTTTTTAAG	TTTCACTCAC	ATAGCCAACT	3750
CTCCCAAAGG	GCACACCCCT	GGGGCTGAGT	CTCCAGGGCC	CCCCAACTGT	3800
GGTAGCTCCA	GCGATGGTGC	TGCCCAGGCC	TCTCGGTGCT	CCATCTCCGC	3850
CTCCACACTG	ACCAAGTGCT	GGCCCACCCA	GTCCATGCTC	CAGGGTCAGG	3900
CGGAGCTGCT	GAGTGACAGC	TTTCCTCAAA	AAGCAGAAGG	AGAGTGAGTG	3950
CCTTTCCCTC	CTAAAGCTGA	ATCCCGGCGG	AAAGCCTCTG	TCCGCCTTTA	4000
CAAGGGAGAA	GACAACAGAA	AGAGGGACAA	GAGGGTTCAC	ACAGCCCAGT	4050
TCCCGTGACG	AGGCTCAAAA	ACTTGATCAC	ATGCTTGAAT	GGAGCTGGTG	4100
AGATCAACAA	CACTACTTCC	CTGCCGGAAT	GAACTGTCCG	TGAATGGTCT	4150
CTGTCAAGCG	GGCCGTCTCC	CTTGGCCCAG	AGACGGAGTG	TGGGAGTGAT	4200
TCCCAACTCC	TTTCTGCAGA	CGTCTGCCTT	GGCATCCTCT	TGAATAGGAA	4250
GATCGTTCCA	CTTTCTACGC	AATTGACAAA	CCCGGAAGAT	CAGATGCAAT	4300
TGCTCCCATC	AGGGAAGAAC	CCTATACTTG	GTTTGCTACC	CTTAGTATTT	4350
ATTACTAACC	TCCCTTAAGC	AGCAACAGCC	TACAAAGAGA	TGCTTGGAGC	4400
AATCAGAACT	TCAGGTGTGA	CTCTAGCAAA	GCTCATCTTT	CTGCCCGGCT	4450
ACATCAGCCT	TCAAGAATCA	GAAGAAAGCC	AAGGTGCTGG	ACTGTTACTG	4500
ACTTGGATCC	CAAAGCAAGG	AGATCATTTG	GAGCTCTTGG	GTCAGAGAAA	4550
ATGAGAAAGG	ACAGAGCCAG	CGGCTCCAAC	TCCTTTCAGC	CACATGCCCC	4600
AGGCTCTCGC	TGCCCTGTGG	ACAGGATGAG	GACAGAGGGC	ACATGAACAG	4650
CTTGCCAGGG	ATGGGCAGCC	CAACAGCACT	TTTCCTCTTC	TAGATGGACC	4700
			*****		47EA

CCAGCATTTA AGTGACCTTC TGATCTTGGG AAAACAGCGT CTTCCTTCTT

TATCTATAGC AACTCATTGG TGGTAGCCAT CAAGCACTTC GGAATT

PCT/US99/11743

4750

4796

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 924 (B) TYPE: protein

WO 99/60986

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: no (vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(ix) FEATURE: Huntingtin-interacting protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Ser Arg Met Trp Gly His Leu Ser Glu Gly Tyr Gly Gln Leu 1 5 10

Cys Ser Ile Tyr Leu Lys Leu Leu Arg Thr Lys Met Glu Tyr His
20 25 30

	WO 99	9/60986											PCT/	US99/1
Thr	Lys	Asn	Pro	Arg 35	Phe	Pro	Gly	Asn	Leu 40	Gln	Met	Ser	Asp	Arg 45
Gln	Leu	Asp	Glu	Ala 50	Gly	Glu	Ser	Asp	Val 55	Asn	Asn	Phe	Phe	Gln 60
Leu	Thr	Val	Glu	Met 65	Phe	Asp	Tyr	Leu	Glu 70	Cys	Glu	Leu	Asn	Leu 75
Phe	Gln	Thr	Val	Phe 80	Asn	Ser	Leu	Asp	Met 85	Ser	Arg	Ser	Val	Ser 90
Val	Thr	Ala	Ala	Gly 95	Gln	Cys	Arg	Leu	Ala 100	Pro	Leu	Ile	Gln	Val 105
Ile	Leu	Asp	Cys	Ser 110	His	Leu	Tyr	Asp	Tyr 115	Thr	Val	Lys	Leu	Leu 120
Phe	Lys	Leu	His	Ser 125	Cys	Leu	Pro	Ala	Asp 130	Thr	Leu	Gln	Gly	His 135
Arg	Asp	Arg	Phe	Met 140	Glu	Gln	Phe	Thr	Lys 145	Leu	Lys	Asp	Leu	Phe 150
Tyr	Arg	Ser	Ser	Asn 155	Leu	Gln	Tyr	Phe	Lys 160	Arg	Leu	Ile	Gln	Ile 165
Pro	Gln	Leu	Pro	Glu 170	Asn	Pro	Pro	Asn	Phe 175	Leu	Arg	Ala	Ser	Ala 180
Leu	Ser	Glu	His	Ile 185	Ser	Pro	Val	Val	Val 190	Ile	Pro	Ala	Glu	Ala 195
Ser	Ser	Pro	Asp	Ser 200	Glu	Pro	Val	Leu	Glu 205	Lys	Asp	Asp	Leu	Met 210
Asp	Met	Asp	Ala	Ser 215	Gln	Gln	Asn	Leu	Phe 220	Asp	Asn	Lys	Phe	Asp 225
Asp	Ile	Phe	Gly	Ser 230	Ser	Phe	Ser	Ser	Asp 235	Pro	Phe	Asn	Phe	Asn 240
Ser	Gln	Asn	Gly	Val 245	Asn	Lys	Asp	Glu	Lys 250	Asp	His	Leu	Ile	Glu 255
Arg	Leu	Tyr	Arg	Glu 260	Ile	Ser	Gly	Leu	Lys 265	Ala	Gln	Leu	Glu	Asn 270
Met	Lys	Thr	Glu	Ser 275	Gln	Arg	Val	Val	Leu 280	Gln	Leu	Lys	Gly	His 285

PC													PCT/I	US99/11743
	WO 99			_										
Val	Ser	Glu	Leu	Glu 290	Ala	Asp	Leu	Ala	Glu 295	Gln	Gln	His	Leu	Arg 300
Gln	Gln	Ala	Ala	Asp 305	Asp	Cys	Glu	Phe	Leu 310	Arg	Ala	Glu	Leu	Asp 315
Glu	Leu	Arg	Arg	Gln 320	Arg	Glu	Asp	Thr	Glu 325	Lys	Ala	Gln	Arg	Ser 330
Leu	Ser	Glu	Ile	Glu 335	Arg	Lys	Ala	Gln	Ala 340	Asn	Glu	Gln	Arg	Tyr 345
Ser	Lys	Leu	Lys	Glu 350	Lys	Tyr	Ser	Glu	Leu 355	Val	Gln	Asn	His	Ala 360
Asp	Leu	Leu	Arg	Lys 365	Asn	Ala	Glu	Val	Thr 370	Lys	Gln	Val	Ser	Met 375
Ala	Arg	Gln	Ala	Gln 380	Val	Asp	Leu	Glu	Arg 385	Glu	Lys	Lys	Glu	Leu 390
Glu	Asp	Ser	Leu	Glu 395	Arg	Ile	Ser	Asp	Gln 400	Gly	Gln	Arg	Lys	Thr 405
Gln	Glu	Gln	Leu	Glu 410	Val	Leu	Glu	Ser	Leu 415	Lys	Gln	Glu	Leu	Gly 420
Thr	Ser	Gln	Arg	Glu 425	Leu	Gln	Val	Leu	Gln 430	Gly	Ser	Leu	Glu	Thr 435
Ser	Ala	Gln	Ser	Glu 440	Ala	Asn	Trp	Ala	Ala 445	Glu	Phe	Ala	Glu	Leu 450
Glu	Lys	Glu	Arg	Asp 455	Ser	Leu	Val	Ser	Gly 460	Ala	Ala	His	Arg	Glu 465
Glu	Glu	Leu	Ser	Ala 470	Leu	Arg	Lys	Glu	Leu 475	Gln	Asp	Thr	Gln	Leu 480
Lys	Leu	Ala	Ser	Thr 485		Glu	. Ser	Met	. Cys 490		Leu	Ala	Lys	Asp 495
Gln	Arg	Lys	Met	Leu 500		. Val	Gly	ser Ser	Arg 505		Ala	Ala	. Glu	Gln 510
Val	Ile	e Glr	n Asp	Ala 515		a Asr	n Glr	ı Lev	Glu 520		Pro) Pro	Leu	Ile 525
Ser	Cys	s Alá	a Gly	Ser 530		a Asp	His	s Lev	Lei 535		Thr	· Val	Thr	Ser 540

	WO 99/60986 PCT/US99/11743													
Ile	Ser	Ser	Cys	Ile 545	Glu	Gln	Leu	Glu	Lys 550	Ser	Trp	Ser	Gln	Tyr 555
Leu	Ala	Cys	Pro	Glu 560	Asp	Ile	Ser	Gly	Leu 565	Leu	His	Ser	Ile	Thr 570
Leu	Leu	Ala	His	Leu 575	Thr	Ser	Asp	Ala	Ile 580	Ala	His	Gly	Ala	Thr 585
Thr	Cys	Leu	Arg	Ala 590	Pro	Pro	Glu	Pro	Ala 595	Asp	Ser	Leu	Thr	Glu 600
Ala	Cys	Lys	Gln	Tyr 605	Gly	Arg	Glu	Thr	Leu 610	Ala	Tyr	Leu	Ala	Ser 615
Leu	Glu	Glu	Glu	Gly 620	Ser	Leu	Glu	Asn	Ala 625	Asp	Ser	Thr	Ala	Met 630
Arg	Asn	Cys	Leu	Ser 635	Lys	Ile	Lys	Ala	Ile 640	Gly	Glu	Glu	Leu	Leu 645
Pro	Arg	Gly	Leu	Asp 650	Ile	Lys	Gln	Glu	Glu 655	Leu	Gly	Asp	Leu	Val 660
Asp	Lys	Glu	Met	Ala 665	Ala	Thr	Ser	Ala	Ala 670	Ile	Glu	Thr	Cys	Thr 675
Ala	Arg	Ile	Glu	Glu 680	Met	Leu	Ser	Lys	Ser 685	Arg	Ala	Gly	Asp	Thr 690
Gly	Val	Lys	Leu	Glu 695	Val	Asn	Glu	Arg	Ile 700	Leu	Arg	Cys	Cys	Thr 705
Ser	Leu	Met	Gln	Ala 710	Ile	Gln	Val	Leu	Ile 715	Val	Ala	Ser	Lys	Asp 720
Leu	Gln	Arg	Glu	Ile 725	Val	Glu	Ser	Gly	Arg 730	Gly	Thr	Ala	Ser	Pro 735
Lys	Glu	Phe	Tyr	Ala 740	Lys	Asn	Ser	Arg	Trp 745	Thr	Glu	Gly	Leu	Ile 750
Ser	Ala	Ser	Lys	Ala 765	Val	Gly	Trp	Gly	Ala 770	Thr	Val	Met	Val	Asp 775
Ala	Ala	Asp	Leu	Val 780	Val	Gln	Gly	Arg	Gly 785	Lys	Phe	Glu	Glu	Leu 790
Met	Val	Cys	Ser	His 795	Glu	Ile	Ala	Ala	Ser 800	Thr	Ala	Gln	Leu	Val 805

	WO 99	/60986	•		,								PCT/	US99/11743
Ala	Ala	Ser	Lys	Val 810	Lys	Ala	Asp	Lys	Asp 815	Ser	Pro	Asn	Leu	Ala 820
Gln	Leu	Gln	Gln	Ala 825	Ser	Arg	Gly	Val	Asn 830	Gln	Ala	Thr	Ala	Gly 835
Val	Val	Ala	Ser	Thr 840	Ile	Ser	Gly	Lys	Ser 845	Gln	Ile	Glu	Glu	Thr 850
Asp	Asn	Met	Asp	Phe 855	Ser	Ser	Met	Thr	Leu 860	Thr	Gln	Ile	Lys	Arg 865
Gln	Glu	Met	Asp	Ser 870	Gln	Val	Arg	Val	Leu 875	Glu	Leu	Glu	Asn	Glu 880
Leu	Gln	Lys	Glu	Arg 885	Gln	Lys	Leu	Gly	Glu 890	Leu	Arg	Lys	Lys	His 895
Tyr	Glu	Leu	Ala	Gly 900	Val	Ala	Glu	Gly	Trp 905	Glu	Glu	Gly	Thr	Glu 910
Ala	Ser	Pro	Pro	Thr 915	Leu	Gln	Glu	Val	Val 920	Thr	Glu	Lys	Glu 924	

- (2) INFORMATION FOR SEQ ID NO: 5
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1090
- (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: Huntingtin-interacting protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met Leu Cys Gln Gly Ser Glu Trp Arg Arg Asp Gln Gln Leu 5 10 15

Gly Thr Ala Asn Ala Arg Gln Trp Cys Pro Leu Pro Gln Asp Ala 20 25 30

Gln Pro Ala Gly Ser Trp Glu Arg Cys Pro Pro Leu Pro Pro Ala 35 40 45

Gly Arg Leu Gln Gly Thr Asp His Pro Trp Gly Trp Gly Arg Leu 50 55 60

WO 99/60986		PCT/US99/11743	•

,	WO 99	/60986								_			r C I/C	1377/11
Ala	Gly	Gly	Gly	Glu 65	Arg	Gly	Gly	Leu	Trp 70	Glu	Gly	Leu	Ser	His 75
Ser	Gln	Arg	Leu	Ile 80	His	Leu	Ile	Leu	Leu 85	Ser	Leu	Pro	Leu	Leu 90
Val	Phe	Gln	Thr	Val 95	Ser	Ile	Asn	Lys	Ala 100	Ile	Asn	Thr	Gln	Glu 105
Val	Ala	Val	Lys	Glu 110	Lys	His	Ala	Arg	Thr 115	Cys	Ile	Leu	Gly	Thr 120
His	His	Glu	Lys	Gly 125	Ala	Gln	Thr	Phe	Trp 130	Ser	Va1	Val	Asn	Arg 135
Leu	Pro	Leu	Ser	Ser 140	Asn	Ala	Val	Leu	Cys 145	Trp	Lys	Phe	Cys	His 150
Val	Phe	His	Lys	Leu 155	Leu	Arg	Asp	Gly	His 160	Pro	Asn	Val	Leu	Lys 165
Asp	Ser	Leu	Arg	Tyr 170	Arg	Asn	Glu	Leu	Ser 175	Asp	Met	Ser	Arg	Met 180
Trp	Gly	His	Leu	Ser 185	Glu	Gly	Tyr	Gly	Gln 190	Leu	Cys	Ser	Ile	Tyr 195
Leu	Lys	Leu	Leu	Arg 200	Thr	Lys	Met	Glu	Tyr 205	His	Thr	Lys	Asn	Pro 210
Arg	Phe	Pro	Gly	Asn 215	Leu	Gln	Met	Ser	Asp 220	Arg	Gln	Leu	Asp	Glu 225
Ala	Gly	Glu	Ser	Asp 230	Val	Asn	Asn	Phe	Phe 235	Gln	Leu	Thr	Val	Glu 240
Met	Phe	Asp	Tyr	Leu 245	Glu	Cys	Glu	Leu	Asn 250	Leu	Phe	Gln	Thr	Val 255
Phe	Asn	Ser	Leu	Asp 260	Met	Ser	Arg	Ser	Val 265	Ser	Val	Thr	Ala	Ala 270
Gly	Gln	Cys	Arg	Leu 275	Ala	Pro	Leu	Ile	Gln 288	Val	Ile	Leu	Asp	Cys 285
Ser	His	Leu	Tyr	Asp 290	Tyr	Thr	Val	Lys	Leu 295	Leu	Phe	Lys	Leu	His 300
Ser	Cys	Leu	Pro	Ala 305		Thr	Leu	Gln	Gly 310		Arg	Asp	Arg	Phe 315

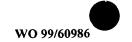
)									
,	WO 99	/60986											PCT/	US99/11743
Met	Glu	Gln	Phe	Thr 320	Lys	Leu	Lys	Asp	Leu 325	Phe	Tyr	Arg	Ser	Ser 330
Asn	Leu	Gln	Tyr	Phe 335	Lys	Arg	Leu	Ile	Gln 340	Ile	Pro	Gln	Leu	Pro 345
Glu	Asn	Pro	Pro	Asn 350	Phe	Leu	Arg	Ala	Ser 355	Ala	Leu	Ser	Glu	His 360
Ile	Ser	Pro	Val	Val 365	Val	Ile	Pro	Ala	Glu 370	Ala	Ser	Ser	Pro	Asp 375
Ser	Glu	Pro	Val	Leu 380	Glu	Lys	Asp	Asp	Leu 385	Met	Asp	Met	Asp	Ala 390
Ser	Gln	Gln	Asn	Leu 395	Phe	Asp	Asn	Lys	Phe 400	Asp	Asp	Ile	Phe	Gly 405
Ser	Ser	Phe	Ser	Ser 410	Asp	Pro	Phe	Asn	Phe 415	Asn	Ser	Gln	Asn	Gly 420
Val	Asn	Lys	Asp	Glu 4 25	Lys	Asp	His	Leu	Ile 430	Glu	Arg	Leu	Tyr	Arg 435
Glu	Ile	Ser	Gly	Leu 440	Lys	Ala	Gln	Leu	Glu 445	Asn	Met	Lys	Thr	Glu 450
Ser	Gln	Arg	Val	Val 455	Leu	Gln	Leu	Lys	Gly 460	His	Val	Ser	Glu	Leu 465
Glu	Ala	Asp	Leu	Ala 470	Glu	Gln	Gln	His	Leu 475	Arg	Gln	Gln	Ala	Ala 480
Asp	Asp	Cys	Glu	Phe 485	Leu	Arg	Ala	Glu	Leu 490	Asp	Glu	Leu	Arg	Arg 495
Gln	Arg	Glu	Asp	Thr 500	Glu	Lys	Ala	Gln	Arg 505	Ser	Leu	Ser	Glu	Ile 510
Glu	Arg	Lys	Ala	Gln 515	Ala	Asn	Glu	Gln	Arg 520	Tyr	Ser	Lys	Leu	Lys 525
Glu	Lys	Tyr	Ser	Glu 530	Leu	Val	Gln	Asn	His 535	Ala	Asp	Leu	Leu	Arg 540
Lys	Asn	Ala	Glu	Val 545	Thr	Lys	Gln	Val	Ser 550	Met	Ala	Arg	Gln	Ala 555
Gln	Val	Asp	Leu	Glu 560	Arg	Glu	Lys	Lys	Glu 565	Leu	Glu	Asp	Ser	Leu 570

	WO 99	/60986											PCT/	US99/11743
Glu	Arg	Ile	Ser	Asp	Gln	Gly	Gln	Arg	Lys	Thr	Gln	Glu	Gln	Leu
				575					588					585

	W O 33	/00200												
Glu	Arg	Ile	Ser	Asp 575	Gln	Gly	Gln	Arg	Lys 588	Thr	Gln	Glu	Gln	Leu 585
Glu	Val	Leu	Glu	Ser 590	Leu	Lys	Gln	Glu	Leu 595	Ala	Thr	Ser	Gln	Arg 600
Glu	Leu	Gln	Val	Leu 605	Gln	Gly	Ser	Leu	Glu 610	Thr	Ser	Ala	Gln	Ser 615
Glu	Ala	Asn	Trp	Ala 620	Ala	Glu	Phe	Ala	Glu 625	Leu	Glu	Lys	Glu	Arg 630
Asp	Ser	Leu	Val	Ser 635	Gly	Ala	Ala	His	Arg 6 4 0	Glu	Glu	Glu	Leu	Ser 645
Ala	Leu	Arg	Lys	Glu 650	Leu	Gln	Asp	Thr	Gln 655	Leu	Lys	Leu	Ala	Ser 660
Thr	Glu	Glu	Ser	Met 665	Cys	Gln	Leu	Ala	Lys 670	Asp	Gln	Arg	Lys	Met 675
Leu	Leu	Val	Gly	Ser 680	Arg	Lys	Ala	Ala	Glu 685	Gln	Val	Ile	Gln	Asp 690
Ala	Leu	Asn	Gln	Leu 695	Gļu	Glu	Pro	Pro	Leu 700	Ile	Ser	Cys	Ala	Gly 705
Ser	Ala	Asp	His	Leu 710	Leu	Ser	Thr	Val	Thr 715	Ser	Ile	Ser	Ser	Cys 720
Ile	Glu	Gln	Leu	Glu 725	Lys	Ser	Trp	Ser	Gln 730	Tyr	Leu	Ala	Cys	Pro 735
Glu	Asp	Ile	Ser	Gly 740	Leu	Leu	His	Ser	Ile 745	Thr	Leu	Leu	Ala	His 750
Leu	Thr	Ser	Asp	Ala 755	Ile	Ala	His	Gly	Ala 760	Thr	Thr	Cys	Leu	Arg 765
Ala	Pro	Pro	Glu	Pro 770	Ala	Asp	Ser	Leu	Thr 775	Glu	Ala	Cys	Lys	Gln 780
Tyr	Gly	Arg	Glu	Thr 785	Leu	Ala	Tyr	Leu	Ala 790	Ser	Leu	Glu	Glu	Glu 795
Gly	Ser	Leu	Glu	Asn 800	Ala	Asp	Ser	Thr	Ala 805	Met	Arg	Asn	Cys	Leu 810
Ser	Lys	Ile	Lys	Ala 815		Gly	Glu	Glu	Leu 820	Leu	Pro	Arg	Gly	Leu 825

WO 99/60986	PCT/US99/11743	
-------------	----------------	--

	WU 99	/60980									_		1 C 1/	03/2/11/43
Asp	Ile	Lys	Gln	Glu 830	Glu	Leu	Gly	Asp	Leu 835	Val	Asp	Lys	Glu	Met 840
Ala	Ala	Thr	Ser	Ala 845	Ala	Ile	Glu	Thr	Ala 850	Thr	Ala	Arg	Ile	Glu 855
Glu	Met	Leu	Ser	Lys 860	Ser	Arg	Ala	Gly	Asp 865	Thr	Gly	Val	Lys	Leu 870
Glu	Val	Asn	Glu	Arg 875	Ile	Leu	Gly	Cys	Cys 888	Thr	Ser	Leu	Met	Gln 885
Ala	Ile	Gln	Val	Leu 890	Ile	Val	Ala	Ser	Lys 895	Asp	Leu	Gln	Arg	Glu 900
Ile	Val	Glu	Ser	Gly 905	Arg	Gly	Thr	Ala	Ser 910	Pro	Lys	Glu	Phe	Tyr 915
Ala	Lys	Asn	Ser	Arg 920	Trp	Thr	Glu	Gly	Leu 925	Ile	Ser	Ala	Ser	Lys 930
Ala	Val	Gly	Trp	Gly 935	Ala	Thr	Val	Met	Val 940	Asp	Ala	Ala	Asp	Leu 945
Val	Val	Gln	Gly	Arg 950	Gly	Lys	Phe	Glu	G1u 955	Leu	Met	Val	Cys	Ser 960
His	Glu	Ile	Ala	Ala 965	Ser	Thr	Ala	Gln	Leu 970	Val	Ala	Ala	Ser	Lys 975
Val	Lys	Ala	Asp	Lys 980	Asp	Ser	Pro	Asn	Leu 985	Ala	Gln	Leu	Gln	Gln 990
Ala	Ser	Arg	Gly	Val 995	Asn	Gln	Ala		Ala 1000	Gly	Val	Val		Ser 1005
Thr	Ile	Ser		Lys 1010	Ser	Gln	Ile		Glu 1015	Thr	Asp	Asn		Asp 1020
Phe	Ser	Ser		Thr 1025	Leu	Thr	Gln		Lys 1030	Arg	Gln	Glu		Asp 1035
Ser	Gln	Val		Val 1040	Leu	Glu	Leu		Asn 1045	Glu	Leu	Gln		Glu 1050
Arg	Gln	Lys		Gly 1055	Glu	Leu	Arg		Lys 1060	His	Tyr	Glu		Ala 1065
Gly	Val	Ala		Gly 1070	Trp	Glu	Glu		Thr 1075	Glu	Ala	Ser		Pro 1080



Thr Leu Gln Glu Val Val Thr Glu Lys Glu 1085 1090

- (2) INFORMATION FOR SEQ ID NO:6:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3301(B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: cDNA for Huntingtin-interacting protein
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGGTGAGCTG	GAGGAGCAGC	GGAAGCAGAA	GCAGAAGGCC	CTGGTGGATA	50
ATGAGCAGCT	CCGCCACGAG	CTGGCCCAGC	TGAGGGCTGC	CCAGCTGGAG	100
CGCGAGCGGA	GCCAGGGCCT	GCGTGAGGAG	GCTGAGAGGA	AGGCCAGTGC	150
CACGGAGGCG	CGCTACAACA	AGCTGAAGGA	AAAGCACAGT	GAGCTCGTCC	200
ATGTGCACGC	GGAGCTGCTC	AGAAAGAACG	CGGACACAGC	CAAGCAGCTG	250
ACGGTGACGC	AGCAAAGCCA	GGAGGAGGTG	GCGCGGTGA	AGGAGCAGCT	300
GGCCTTCCAG	GTGGAGCAGG	TGAAGCGGGA	GTCGGAGTTG	AAGCTAGAGG	350
AGAAGAGCGA	CCAGCAGGAG	AAGCTCAAGA	GGGAGCTGGA	GGCCAAGGCC	400
GGAGAGCTGG	CCCGCGCGCA	GGAGGCCCTG	AGCCACACAG	AGCAGAGCAA	450
GTCGGAGCTG	AGCTCACGGC	TGGACACACT	GAGTGCGGAG	AAGGATGCTC	500
TGAGTGGAGC	TGTGCGGCAG	CGGGAGGCAG	ACCTGCTGGC	GGCGCAGAGC	550
CTGGTGCGCG	AGACAGAGGC	GGCGCTGAGC	CGGGAGCAGC	AGCGCAGCTC	600
CCAGGAGCAG	GGCGAGTTGC	AGGGCCGGCT	GGCAGAGAGG	GAGTCTCAGG	650
AGCAGGGGCT	GCGGCAGAGG	CTGCTGGACG	AGCAGTTCGC	AGTGTTGCGG	700
GGCGCTGCTG	CCGAGGCCGC	GGGCATCCTG	CAGGATGCCG	TGAGCAAGCT	750
GGACGACCCC	CTGCACCTGC	GCTGTACCAG	CTCCCCAGAC	TACCTGGTGA	800
GCAGGGCCCA	GGAGGCCTTG	GATGCCGTGA	GCACCCTGGA	GGAGGGCCAC	850
GCCCAGTACC	TGACCTCCTT	GGCAGACGCC	TCCGCCCTGG	TGGCAGCTCT	900
GACCCGCTTC	TCCCACCTGG	CTGCGGATAC		GGCGGTGCCA	950
CCTCGCACCT	GGCTCCCACC	GACCCTGCCG	ACCGCCTCAT	AGACACCTGC	1000
AGGGAGTGCG	GGGCCCGGGC	TCTGGAGCTC	ATGGGGCAGC	TGCAGGACCA	1050
GCAGGCTCTG	CGGCACATGC	AGGCCAGCCT	GGTGCGGACA	CCCCTGCAGG	1 1 00
GCATCCTTCA	GCTGGGCCAA	GAACTGAAAC	CCAAGAGCCT	AGATGTGCGG	1150
CAGGAGGAGC	TGGGGGCCGT	GGTCGACAAG	GAGATGGCGG	CCACATCCGC	1200
AGCCATTGAA	GATGCTGTGC	GGAGGATTGA	GGACATGATG	AACCAGGCAC	1250
GCCACGCCAG	CTCGGGGGTG	AAGCTGGAGG	TGAACGAGAG	GATCCTCAAC	1300
TCCTGCACAG	ACCTGATGAA	GGCTATCCGG	CTCCTGGTGA		1350
TAGCCTGCAG	AAGGAGATCG	TGGAGAGCGG	CAGGGGGGCA	GCCACGCAGC	1400
AGGAATTTTA	CGCCAAGAAC	TCGCGCTGGA	CCGAAGGCCT	CATCTCGGCC	1450
TCCAAGGCTG	TGGGCTGGGG			CAGCTGACAA	1500
GGTGGTGCTT			GCTCATCGTC	TGCTCCCACG	1550
AGATCGCAGC	CAGCACGGCC	CAGCTGGTGG	CGGCCTCCAA	GGTGAAGGCC	1600

WO 99/60986 PCT/US99/11743

AACAAGCACA	GCCCCCACCT	GAGCCGCCTG	CAGGAATGTT	CTCGCACAGT	1650
CAATGAGAGG	GCTGCCAATG	TGGTGGCCTC	CACCAAGTCA	GGCCAGGAGC	1700
AGATTGAGGA	CAGAGACACC	ATGGATTTCT	CCGGCCTGTC	CCTCATCAAG	1750
CTGAAGAAGC	AGGAGATGGA	GACGCAGGTG	CGTGTCCTGG	AGCTGGAGAA	1800
GACGCTGGAG	GCTGAACGCA	TGCGGCTGGG	GGAGTTGCGG	AAGCAACACT	1850
ACGTGCTGGC	TGGGGCATCA	GGCAGCCCTG	GAGAGGAGGT	GGCCATCCGG	1900
CCCAGCACTG	CCCCCGAAG	TGTAACCACC	AAGAAACCAC	CCCTGGCCCA	1950
GAAGCCCAGC	GTGGCCCCCA	GACAGGACCA	CCAGCTTGAC	AAAAAGGATG	2000
GCATCTACCC	AGCTCAACTC	GTGAACTACT	AGGCCCCCCA	GGGGTCCAGC	2050
AGGGTGGCTG	GTGACAGGCC	TGGGCCTCTG	CAACTGCCCT	GACAGGACCG	2100
AGAGGCCTTG	CCCCTCCACC	TGGTGCCCAA	GCCTCCCGCC	CCACCGTCTG	2150
GATCAATGTC	CTCAAGGCCC	CTGGCCCTTA	CTGAGCCTGC	AGGGTCCTGG	2200
GCCATGTGGG	TGGTGCTTCT	GGATGTGAGT	CTCTTATTTA	TCTGCAGAAG	2250
GAACTTTGGG	GTGCAGCCAG	GACCCGGTAG	GCCTGAGCCT	CAACTCTTCA	2300
GAAAATAGTG	TTTTTAATAT	TCCTCTTCAG	AAAATAGTGT	TTTTAATATT	2350
CCGAGCTAGA	GCTCTTCTTC	CTACGTTTGT	AGTCAGCACA	CTGGGAAACC	2400
GGGCCAGCGT	GGGGCTCCCT	GCCTTCTGGA	CTCCTGAAGG	TCGTGGATGG	2450
ATGGAAGGCA	CACAGCCCGT	GCCGGCTGAT	GGGACGAGGG	TCAGGCATCC	2500
TGTCTGTGGC	CTTCTGGGGC	ACCGATTCTA	CCAGGCCCTC	CAGCTGCGTG	2550
GTCTCCGCAG	ACCAGGCTCT	GTGTGGGCTA	GAGGAATGTC	GCCCATTACC	2600
TCCTCAGGCC	CTGGCCCTCG	GGCCTCCGTG	ATGGGAGCCC	CCCAGGAGGG	2700
GTCAGATGCT	GGAAGGGGCC	GCTTTCTGGG	GAGTGAGGTG	AGACATAGCG	2750
GCCCAGGCGC	TGCCTTCACT	CCTGGAGTTT	CCATTTCCAG	CTGGAATCTG	2800
CAGCCACCCC	CATTTCCTGT	TTTCCATTCC	CCCGTTCTGG	CCGCGCCCCA	2850
CTGCCCACCT	GAAGGGGTGG	TTTCCAGCCC	TCCGGAGAGT	GGGCTTGGCC	2900
CTAGGCCCTC	CAGCTCAGCC	AGAAAAAGCC	CAGAAACCCA	GGTGCTGGAC	2950
CAGGGCCCTC	AGGGAGGGAC	CCTGCGGCTA	GAGTGGGCTA	GGCCCTGGCT	3000
TTGCCCGTCA	GATTTGAACG	AATGTGTGTC	CCTTGAGCCC	AAGGAGAGCG	3050
GCAGGAGGGG	TGGGACCAGG	CTGGGAGGAC	AGAGCCAGCA	GCTGCCATGC	3100
CCTCCTGCTC	CCCCCACCCC	AGCCCTAGCC	CTTTAGCCTT	TCACCCTGTG	3150
CTCTGGAAAG	GCTACCAAAT	ACTGGCCAAG	GTCAGGAGGA	GCAAAAATGA	3200
GCCAGCACCA	GCGCCTTGGC	TTTGTGTTAG		TGAAGTGTTC	3250
TGTTGGCAAT	AAAATGCACT	TTGACTGTTA	AAAAAAAAA	AAAAAAAAA	3300
A					3301

- (2) INFORMATION FOR SEQ ID NO: 7
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 676 (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE: (A) ORGANISM: human
- (ix) FEATURE: Huntingtin-interacting protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
- Gly Glu Leu Glu Glu Gln Arg Lys Gln Lys Gln Lys Ala Leu Val

5 10 15

WO 99/6098	WO 99/60986								PCT/	US99/1174	43
	~ 7	_	_			_	 		_	_	

	WO 99	/60986							`				PCT/	US99/11
Asp	Asn	Glu	Gln	Leu 20	Arg	His	Glu	Leu	Ala 25	Gln	Leu	Arg	Ala	Ala 30
Gln	Leu	Glu	Arg	Glu 35	Arg	Ser	Gln	Gly	Leu 40	Arg	Glu	Glu	Ala	Glu 45
Arg	Lys	Ala	Ser	Ala 50	Thr	Glu	Ala	Arg	Tyr 55	Asn	Lys	Leu	Lys	Glu 60
Lys	His	Ser	Glu	Leu 65	Val	His	Val	His	Ala 70	Glu	Leu	Leu	Arg	Lys 75
Asn	Ala	Asp	Thr	Ala 80	Lys	Gln	Leu	Thr	Val 85	Thr	Gln	Gln	Ser	Gln 90
Glu	Glu	Val	Ala	Arg 95	Val	Lys	Glu	Gln	Leu 100	Ala	Phe	Gln	Val	Glu 105
Gln	Val	Lys	Arg	Glu 110	Ser	Glu	Leu	Lys	Leu 115	Glu	Glu	Lys	Ser	Asp 120
Gln	Gln	Glu	Lys	Leu 125	Lys	Arg	Glu	Leu	Glu 130	Ala	Lys	Ala	Gly	Glu 135
Leu	Ala	Arg	Ala	Gln 140	Glu	Ala	Leu	Ser	His 145	Thr	Glu	Gln	Ser	Lys 150
Ser	Glu	Leu	Ser	Ser 155	Arg	Leu	Asp	Thr	Leu 160	Ser	Ala	Glu	Lys	Asp 165
Ala	Leu	Ser	Gly	Ala 170	Val	Arg	Gln	Arg	Glu 175	Ala	Asp	Leu	Leu	Ala 180
Ala	Gln	Ser	Leu	Val 185	Arg	Glu	Thr	Glu	Ala 190	Ala	Leu	Ser	Arg	Glu 195
Gln	Gln	Arg	Ser	Ser 200	Gln	Glu	Gln	Gly	Glu 205	Leu	Gln	Gly	Arg	Leu 210
Ala	Glu	Arg	Glu	Ser 215	Gln	Glu	Gln	Gly	Leu 220	Arg	Gln	Arg	Leu	Leu 225
Asp	Glu	Gln	Phe	Ala 230	Val	Leu	Arg	Gly	Ala 235	Ala	Ala	Glu	Ala	Ala 240
Gly	Ile	Leu	Gln	Asp 245	Ala	Val	Ser	Lys	Leu 250	Asp	Asp	Pro	Leu	His 255
Leu	Arg	Cys	Thr	Ser 260	Ser	Pro	Asp	Tyr	Leu 265	Val	Ser	Arg	Ala	Gln 270

)								РСТ /Г	IS99/11743
	VO 99/		7	7.7.	17-1	Com	mh ×	Lon	Clu	C1,,	Clu	Hic		
Glu .	А⊥а	ьeu	Asp	275	vai	ser	1111	цеu	288	Giu	GIY	111.5		285
Tyr	Leu	Thr	Ser	Leu 290	Ala	Asp	Ala	Ser	Ala 295	Leu	Val	Ala	Ala	Leu 300
Thr	Arg	Phe	Ser	His 305	Leu	Ala	Ala	Asp	Thr 310	Ile	Ile	Asn	Gly	Gly 315
Ala	Thr	Ser	His	Leu 320	Ala	Pro	Thr	Asp	Pro 325	Ala	Asp	Arg	Leu	Ile 330
Asp	Thr	Cys	Arg	Glu 335	Cys	Gly	Ala	Arg	Ala 340	Leu	Glu	Leu	Met	Gly 345
Gln	Leu	Gln	Asp	Gln 350	Gln	Ala	Leu	Arg	His 355	Met	Gln	Ala	Ser	Leu 360
Val	Arg	Thr	Pro	Leu 365	Gln	Gly	Ile	Leu	Gln 370	Leu	Gly	Gln	Glu	Leu 375
Lys	Pro	Lys	Ser	Leu 380	Asp	Val	Arg	Gln	Glu 385	Glu	Leu	Gly	Ala	Val 390
Val	Asp	Lys	Glu	Met 395	Ala	Ala	Thr	Ser	Ala 400	Ala	Ile	Glu	Asp	Ala 405
Val	Arg	Arg	Ile	Glu 410	Asp	Met	Met	Asn	Gln 415	Ala	Arg	His	Ala	Ser 420
Ser	Gly	Val	Lys	Leu 425	Glu	Val	Asn	Glu	Arg 430	Ile	Leu	Asn	Ser	Cys 435
Thr	Asp	Leu	Met	Lys 440	Ala	Ile	Arg	Leu	Leu 445	Val	Thr	Thr	Ser	Thr 450
Ser	Leu	Gln	Lys	Glu 455	Ile	Val	Glu	Ser	Gly 460	Arg	Gly	Ala	Ala	Thr 465
Gln	Gln	Glu	Phe	Туr 470	Ala	Lys	Asn	Ser	Arg 475	Trp	Thr	Glu	Gly	Leu 480
Ile	Ser	Ala	Ser	Lys 485	Ala	Val	Gly	Trp	Gly 490	Ala	Thr	Gln	Leu	Val 495
Glu	Ala	Ala	Asp	Lys 500		Val	Leu	His	Thr 505	Gly	Lys	Tyr	Glu	Glu 510
Leu	Ile	· Val	. Cys	Ser 515		Glu	Ile	Ala	Ala 520		Thr	Ala	Gln	Leu 525

WO 99/60986			PCT/US99/11743							
Val Ala Ala Ser	Lys Val Lys	Ala Asn Lys His	Ser Pro His Leu							
	530	535	540							
Ser Arg Leu Glr	Glu Cys Ser 545	Arg Thr Val Asn 550	Glu Arg Ala Ala 555							
Asn Val Val Ala	Ser Thr Lys	Ser Gly Gln Glu	Gln Ile Glu Asp							
	560	565	570							
Arg Asp Thr Met	Asp Phe Ser	Gly Leu Ser Leu	Ile Lys Leu Lys							
	575	588	585							
Lys Gln Glu Met	Glu Thr Gln	Val Arg Val Leu	Glu Leu Glu Lys							
	590	595	600							
Thr Leu Glu Ala	Glu Arg Met	Arg Leu Gly Glu	Leu Arg Lys Gln							
	605	610	615							
His Tyr Val Leu	Ala Gly Ala	Ser Gly Ser Pro	Gly Glu Glu Val							
	620	625	630							
Ala Ile Arg Pro	Ser Thr Ala 635	Pro Arg Ser Val 640	Thr Thr Lys Lys 645							
Pro Pro Leu Ala	Gln Lys Pro	Ser Val Ala Pro	Arg Gln Asp His							
	650	655	660							
Gln Leu Asp Lys	Lys Asp Gly	Ile Tyr Pro Ala	Gln Leu Val Asn							
	665	670	675							
Tyr										
(2) INFORMATION FOR SEQ ID NO:8: (i) SEQUENCE CHARACTERISTICS:										

- (A) LENGTH: 2338(B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: mouse
- (ix) FEATURE: cDNA for Huntingtin-interacting protein mHIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGCACGAGGG	CTCATTCAGA	TCCCCCAGCT	GCCCGAGAAT	CCACCCAACTT	50
CCTACGAGCC	TCGGCCCTGT	CAGAGCACAT	CAGTCCTGTG	GTGGTGATCCC	100
GGCAGAGGTG	TCATCCCCAG	ACAGTGAGCC	TGTCCTGGAG	AAGGATGACCT	150
CATGGACATG	GACGCCTCCC	ACCAGACTTT	CTTTCACAAC	$\Delta \Delta C T T T C \Delta T C \Delta$	200



CCMCMMMCCC	አርርጥር አጥጥር አ	CCACCCACCC	መመመር አ አመመመር	AACAATCAAAA	250
				CTGTACAGAGA	
				GAGAGCCAGCG	
0 0	CAGCTGAAGG			GCAGAGCTAGC	
		-	GGATGACTGC		
	GATGAACTGA		AGAGGACACG	GAGAAGGCACA	
		AAAGAAAGGC	CCAGGCTAAT		
			GGTGCAGAAC	CATGCTGACCT	
GCTGCGGAAG			GGTGTCCGTG	GCCCGGCAAGC	
	TTGGAAAGAG	AGAAAAAAGA		TCCTTTGCAC	700
00110010011	1100.000		AAGAGCAACA	GGATGTTCTA	750
•		GGCCACCAGC		TGCAGGTCCT	800
CCACAGCAAC			AGAAGCGAAA	TGGCTGACAC	850
	GTTGGAGAAG			TGTTGCAGCT	900
			GACCAGCTGG		950
			GTGCCAGCAG		1000
AGAGGAAAAC			AGGCTGCGGA		1050
CAGGAGGCGC			ACCCTCATCA		1100
ATCCACAGAT	CACCTTCTCT		CTCCGTTTCC		1150
	AAAGAACGGC			AGAAGATATT	1200
AGTGAGCTTC		CACCCTGCTT	GCCCACTTGA		1250
			GGCCCCACCG		1300
ACTCGTTGAC			GCAGAGAAAC		1350
CTGTCCTCCC			GAGAATGCTG		1400
CCTTAGGAAT		GGGTCAAGAC		GAGCTGCTGC	1450
CCAGGGGCCT	GGACATCAAG			GGTGGACAAG	1500
GAGATGGCAG	CCACTTCAGC			CCCGGATAGA	
GGAAATTCTC	AGTAAGTCCC		CACGGGAGTC		1600
TGAATGAGAG	GATCCTGGGT		GCCTGATGCA		1650
GTGCTCGTTG			AAGGAGATAG	TGGAGAGTGG	1700
CAGGGGTAGT			CGCCAAGAAC	TCTCGGTGGA	
CGGAAGGGCT		TCCAAAGCTG		AGCTACCATC	1800
ATGGTGGATG	CTGCTGATCT	- -	GGCAAAGGGA		1850
GCTGATGGTG	TGTTCACGCG		CAGTACTGCC		1900
			GCCTCAATCT		2000
				TGGTGGCCTC	
				ATGGACTTCT	
				TTCCCAGGTT	
				AGAAACTAGG	
				GAGGGCTGGG	
				AATACCGGAC	
				MAIACCGGAC	2338
AAAGAGTAGA	GCCAAGCCGA	CACCCCACAC	ATCAGAAA		433 0

- (2) INFORMATION FOR SEQ ID NO: 9:
- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 676(B) TYPE: protein

WO 99/60986

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: protein

- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: mouse
- (ix) FEATURE: Huntingtin-interacting protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9: Ala Arg Gly Leu Ile Gln Ile Pro Gln Leu Pro Glu Asn Pro Pro Asn Phe Leu Arg Ala Ser Ala Leu Ser Glu His Ile Ser Pro Val Val Val Ile Pro Ala Glu Val Ser Ser Pro Asp Ser Glu Pro Val Leu Glu Lys Asp Asp Leu Met Asp Met Asp Ala Ser Gln Gln Thr Leu Phe Asp Asn Lys Phe Asp Asp Val Phe Gly Ser Ser Leu Ser Ser Asp Pro Phe Asn Phe Asn Asn Gln Asn Gly Val Asn Lys Asp Glu Lys Asp His Leu Ile Glu Arg Leu Tyr Arg Glu Ile Ser Gly Leu Thr Gly Gln Leu Asp Asn Met Lys Ile Glu Ser Gln Arg Ala Met Leu Gln Leu Lys Gly Arg Val Ser Glu Leu Glu Ala Glu Leu Ala Glu Gln Gln His Leu Gly Arg Gln Ala Met Asp Asp Cys Glu Phe Leu Arg Thr Glu Leu Asp Glu Leu Lys Arg Gln Arg Glu Asp Thr Glu Lys Ala Gln Arg Ser Leu Thr Glu Ile Glu Arg Lys Ala Gln Ala Asn Glu Gln Arg Tyr Ser Lys Leu Lys Glu Lys Tyr Ser Glu Leu Val Gln Asn His Ala Asp Leu Leu Arg Lys Asn Ala Glu Val Thr Lys Gln Val Ser Val Ala Arg Gln Ala Gln Val Asp Leu

Glu Arg Glu Lys Lys Glu Leu Ala Asp Ser Phe Ala Arg Val Ser

•	WO 99/	60986			,								PCT/U	J S99/11
				230					235					240
Asp	Gln	Ala	Gln	Arg 245	Lys	Thr	Gln	Glu	Gln 250	Gln	Asp	Val	Leu	Glu 255
Asn	Leu	Lys	His	Glu 260	Leu	Ala	Thr	Ser	Arg 265	Gln	Glu	Leu	Gln	Val 270
Leu	His	Ser	Asn	Leu 275	Glu	Thr	Ser	Ala	Gln 288	Ser	Glu	Ala	Lys	Trp 285
Leu	Thr	Gln	Ile	Ala 290	Glu	Leu	Glu	Lys	Glu 295	Gln	Gly	Ser	Leu	Ala 300
Thr	Val	Ala	Ala	Gln 305	Arg	Glu	Glu	Glu	Leu 310	Ser	Ala	Leu	Arg	Asp 315
Gln	Leu	Glu	Ser	Thr 320	Gln	Ile	Lys	Leu	Ala 325	Gly	Ala	Gln	Glu	Ser 330
Met	Cys	Gln	Gln	Val 335	Lys	Asp	Gln	Arg	Lys 340	Thr	Leu	Leu	Ala	Gly 345
Ile	Arg	Lys	Ala	Ala 350	Glu	Arg	Glu	Ile	Gln 355	Glu	Ala	Leu	Ser	Gln 360
Leu	Glu	Glu	Pro	Thr 365	Leu	Ile	Ser	Cys	Ala 370	Gly	Ser	Thr	Asp	His 375
Leu	Leu	Ser	Lys	Val 380	Ser	Ser	Val	Ser	Ser 385	Cys	Leu	Glu	Gln	Leu 390
Glu	Lys	Asn	Gly	Ser 395	Gln	Tyr	Leu		Cys 400	Pro	Glu	Asp	Ile	Ser 405
Glu	Leu	Leu	His	Ser 410	Ile	Thr	Leu	Leu	Ala 415	His	Leu	Thr	Gly	Asp 420
Thr	Val	Ile	Gln	Gly 425	Ser	Ala	Thr	Ser	Leu 430	Arg	Ala	Pro	Pro	Glu 435
Pro	Ala	Asp	Ser	Leu 440	Thr	Glu	Ala	Cys	Arg 445	Gln	Tyr	Gly	Arg	Glu 450
Thr	Leu	Ala	Tyr	Leu 455	Ser	Ser	Leu	Glu	Glu 460	Glu	Gly	Thr	Val	Glu 465
Asn	Ala	Asp	Val	Thr 470	Ala	Leu	Arg	Asn	Cys 475	Leu	Ser	Arg	Val	Lys 480

									4					
,	WO 99/	60986							•				PCT/U	JS99/11743
Thr	Leu	Gly	Glu	Glu 485	Leu	Leu	Pro	Arg	Gly 490	Leu	Asp	Ile	Lys	Gln 495
Glu	Glu	Leu	Gly	Asp 500	Leu	Val	Asp	Lys	Glu 505	Met	Ala	Ala	Thr	Ser 510
Ala	Ala	Ile	Glu	Ala 515	Ala	Thr	Thr	Arg	Ile 520	Glu	Glu	Ile	Leu	Ser 525
Lys	Ser	Arg	Ala	Gly 530	Asp	Thr	Gly	Val	Lys 535	Leu	Glu	Val	Asn	Glu 540
Arg	Ile	Leu	Gly	Ser 545	Cys	Thr	Ser	Leu	Met 550	Gln	Ala	Ile	Lys	Val 555
Leu	Val	Val	Ala	Ser 560	Lys	Asp	Leu	Gln	Lys 565	Glu	Ile	Val	Glu	Ser 570
Gly	Arg	Gly	Ser	Ala 575	Ser	Pro	Lys	Glu	Phe 588	Tyr	Ala	Lys	Asn	Ser 585
Arg	Trp	Thr	Glu	Gly 590	Leu	Ile	Ser	Ala	Ser 595	Lys	Ala	Val	Gly	Trp 600
Gly	Ala	Thr	Ile	Met 605	Val	Asp	Ala	Ala	Asp 610	Leu	Val	Val	Gln	Gly 615
Lys	Gly	Lys	Phe	Glu 620	Glu	Leu	Met	Val	Cys 625	Ser	Arg	Glu	Ile	Ala 630
Ala	Ser	Thr	Ala	Gln 635	Leu	Val	Ala	Ala	Ser 640		Val	Lys	Ala	Asn 645
Lys	Gly	Ser	Leu	Asn 650	Leu	Thr	Gln	Leu	Gln 655		Ala	Ser	Arg	Gly 660
Val	Asn	Gln	Ala	Thr 665	Ala	Ala	Val	Val	Ala 670		Thr	Ile	Ser	Gly 675
Lys	Ser	Gln	Ile	Glu 680		Thr	Asp	Ser	Met 685		Phe	Ser	Ser	Met 690
Thr	Leu	Thr	Gln	Ile 695		Arg	Gln	Glu	Met 700		Ser	Gln	Val	Arg 705
Val	Leu	Glu	Leu	Glu 710		Asp	Leu	Gln	Lys 715		Arg	Gln	Lys	Leu 720
Gly	Glu	. Lev	Arg	1 Lys 725		His	з Туг	Glu	1 Leu 730		Gly	Val	Ala	Glu 735



Gly Trp Glu Glu Gly Thr Glu Ala Ser Pro Ser Thr Val Gln Glu 740 745 750

Ala Ile Pro Asp Lys Glu 755

- (2) INFORMATION FOR SEQ ID NO:10:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3964
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: mouse
- (ix) FEATURE: cDNA for Huntingtin-interacting protein mHIP1a
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGCACGAGGC	GGCGCGCGC	CTCCGTGTGC	CTAGGCTTGA	GGCGGGCGGT	50
GACGCCTCAT	TCGCGCGGAG	CCGGGCCGGG	ACACGGTCGG	CGGCAGCATG	100
AACAGCATCA	AGAATGTGCC	GGCGCGGGTG	CTGAGCCGCA	GGCCGGGCCA	150
CAGCCTAGAG	GCCGAGCGCG	AGCAGTTCGA	CAAGACGCAG	GCCATCAGTA	200
TCAGCAAAGC	CATCAACAGC	CAGGAGGCCC	CAGTGAAGGA	GAAGCATGCC	250
CGGCGTATCA	TCCTGGGCAC	GCATCATGAG	AAGGGAGCCT	TCACCTTCTG	300
GTCCTATGCC	ATCGGCCTGC	CGCTGTCCAG	CAGCTCCATC	CTCAGCTGGA	350
AGTTCTGTCA	CGTCCTTCAC	AAGGTCCTCC	GGGACGGACA	CCCCAACGTC	400
CTGCATGACT	ATCAGCGGTA	CCGGAGCAAC	ATACGTGAGA	TCGGTGACTT	450
GTGGGGCCAC	CTTCGTGACC	AGTATGGACA	CCTGGTGAAT	ATCTATACCA	500
AACTGTTGCT	GACTAAGATC	TCCTTCCACC	TTAAGCACCC	CCAGTTTCCT	550
GCAGGCCTGG	AGGTAACAGA	TGAGGTGTTG	GAGAAGGCGG	CGGGAACTGA	600
TGTCAACAAC	ATTTTTCAGC	TTACCGTGGA	GATGTTTGAC	TACATGGACT	650
GTGAACTGAA	GCTTTCTGAG	TCAGTTTTCC	GGCAGCTCAA	CACGGCCATC	700
GCAGTGTCCC	AGATGTCTTC	TGGCCAGTGT	CGCCTAGCGC	CGCTCATCCA	750
GGTCATTCAG	GACTGCAGCC	ACCTGTACCA	CTACACAGTG	AAGCTCATGT	800
TTAAGCTGCA	CTCCTGTCTC	CCGGCAGACA	CCCTGCAAGG	CCACAGGGAT	850
CGGTTCCACG	AGCAGTTCCA	CAGCCTCAAA	AACTTCTTCC	GCCGGGCTTC	900
AGACATGCTG	TACTTCAAGA	GGCTCATCCA	GATCCCGCGG	CTGCCTGAGG	950
GACCCCCCAA	TTTCCTGCGG	GCTTCAGCCC	TGGCTGAGCA	CATCAAGCCG	1000
GTGGTGGTGA	TTCCCGAGGA	GGCCCCAGAG	GAAGAGGAGC	CTGAGAACCT	1050
AATTGAAATC	AGCAGTGCGC	CCCCTGCTGG	GGAGCCAGTG	GTGGTGGCTG	1100
ACCTCTTTGA	TCAGACCTTT	GGACCCCCCA	ATGGCTCCAT	GAAGGATGAC	1150
AGGGACCTCC	AAATCGAGAA	CTTGAAGAGA	GAGGTGGAGA	CCCTCCGTGC	1200
TGAGCTGGAG	AAGATTAAGA	TGGAGGCACA	GCGGTACATC	TCCCAGCTGA	1250
AGGGCCAGGT	GAATGGCCTG	GAGGCAGAGC	TGGAGGAGCA	GCGCAAGCAG	1300
AAGCAGAAGG	CCCTGGTGGA	CAACGAGCAG	CTGCGCCACG	AGCTGGCCCA	1350
GCTCAAGGCC	CTGCAGCTGG	AGGGCGCCCG	CAACCAGGGC	CTTCGAGAGG	1400
AAGCAGAGAG	GAAGGCCAGT	GCCACGGAGG	CACGCTACAG	CAAGCTGAAG	1450

GAGAAACACA GCGAACTCAT TAACACGCAC GCCGAGCTGC TCAGGAAGAA 1500

WO 99/60986 PCT/US99/11743

			ACAGCAGAGC		1550
			AGATGGAGCA		1600
			GACCAGTTGG		1650
			GGCCCGTGCG		1700
			TGAGCTCACG		1750
			GTCGTTCGGC		1800
			GGAGAAGGAG		1850
			AGGGCGAGCT		1900
			CTTCGGCAGA		1950
			CGCCGAGGCA		2000
			CCCTGCACCT		2050
			CAGGCAGCCC		2100
			CCTGGCTTCC		2150
			TCTCCCATTT		2200
			CTGGCCCCCA		2250
CGACCGCCTG	ATGGACACAT	GCAGGGAGTG	TGGAGCCCGG	GCTCTGGAGC	2300
TGGTGGGACA	GCTGCAAGAC	CAGACAGTGC	TACGGAGGGC	TCAGCCCAGC	2350
CTGATGCGGG	CCCCCTGCA	GGGCATTCTG	CAGTTGGGCC	AGGACTTGAA	2400
GCCTAAGAGC	CTGGATGTAC	GGCAAGAGGA	GCTAGGGGCC	ATGGTGGACA	2450
AGGAGATGGC	GGCCACCTCG	GCAGCCATTG	AGGACGCTGT	GCGGAGGATC	2500
GAGGACATGA	TGAGCCAGGC	CCGCCACGAG	AGCTCAGGCG	TGAAACTGGA	2550
GGTGAATGAG	AGGATCCTCA	ACTCCTGCAC	AGACCTGATG	AAGGCTATCC	2600
GGCTCCTGGT	GATGACCTCC	ACCAGCCTGC	AGAAGGAAAT	TGTGGAGAGC	2650
GGCAGGGGG	CAGCAACGCA	GCAGGAATTT	TATGCCAAGA	ATTCACGGTG	2700
GACTGAAGGC	CTCATCTCAG	CCTCTAAGGC	AGTGGGCTGG	GGAGCCACAC	2750
AGCTGGTGGA	GTCAGCTGAC	AAGGTTGTGC	TTCACATGGG	CAAATACGAG	2800
GAACTCATCG	TCTGCTCCCA	TGAGATTGCG	GCCAGCACGG	CCCAGCTGGT	2850
GGCAGCCTCG	AAGGTGAAAG	CCAACAAGAA	CAGTCCCCAC	TTGAGCCGCC	2900
TGCAGGAATG	TTCCCGCACT	GTCAACGAGA	GGGCTGCCAA	CGTCGTGGCC	2950
TCCACCAAAT	CTGGCCAGGA	GCAGATTGAG	GACAGAGACA	CCATGGATTT	3000
CTCTGGCCTG	TCCCTCATCA	AGTTGAAGAA	GCAGGAGATG	GAGACACAGG	3050
TGCGAGTCTT	GGAGCTGGAG	AAGACACTAG	AGGCAGAGCG	TGTCCGGCTC	3100
GGGGAGCTTC	GGAAACAGCA	CTATGTACTG	GCTGGGGGGA	TGGGAACACC	3150
TAGCGAAGAA	GAACCCAGCA	GACCCAGCCC	AGCTCCCCGA	AGTGGGGCCA	3200
CTAAGAAGCC	ACCGCTGGCC	CAGAAACCCA	GCATAGCCCC	CAGGACAGAC	3250
AACCAGCTCG	A CAAAAAGGA	T GGTGTCTAC	C CAGCTCAAC	T TGTGAACTAC	3300
TAGGCCCCTA	A GGTGTTCAG	C AGGATGGCT	G GTGGTTGTG	C CTGGGCTTCA	3350
TGTGGCTGTC	T GGCAGTGGT	C AAGGGGCCT	C TGAGAAGCC'	r ccaactcctg	3400
CCCAAGGGGC	C TAGTCTGTG	G GACAGTTCA	T CTGGATGTG	A ATCTATTAT	3450
CTTAAGTAGG	A ACTGCCTCG	A GCAGCTGGG	A CCCAGCAGG	C CTGAGCCACA	3500
AATCTGCAGC	G GACATCAGA	G ATAGTCTGA	A TGCTGCGAG	G TATTTCTTTC	3550
TTCGTAAGTT	T AGTCAGCAC	A CTGGGAAAA	G GTCACATAA	G CCAGGAGCCT	3600
CCTTGTCTCT	G GACTCAAAA	G TCTGAGGCC	T TAAGTGAAC	A ACAGAAAGAG	3650
GGTCCCTGCT	G GCTACCAGG	G ATAAGGGGA	T GACCTGTGA	CCTTGAGCCA	3700
				C TGGTGCTAGG	3750
				G GAGCCTGGCA	
				G CCCGTGACCT	
				C TACTAGTGTG	
				C TAAAGCTGGG	
GCCTTTCCTC					3964
5552115516					

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 676 (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: mouse
- (ix) FEATURE: Huntingtin-interacting protein -mHIP1a
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met	Asn	Ser	Ile	Lys 5	Asn	Val	Pro	Ala	Arg 10	Val	Leu	Ser	Arg	Arg 15
Pro	Gly	His	Ser	Leu 20	Glu	Ala	Glu	Arg	Glu 25	Gln	Phe	Asp	Lys	Thr 30
Gln	Ala	Ile	Ser	Ile 35	Ser	Lys	Ala	Ile	Asn 40	Ser	Gln	Glu	Ala	Pro 45
Val	Lys	Glu	Lys	His 50	Ala	Arg	Arg	Ile	Ile 55	Leu	Gly	Thr	His	His 60
Glu	Lys	Gly	Ala	Phe 65	Thr	Phe	Trp	Ser	Tyr 70	Ala	Ile	Gly	Leu	Pro 75
Leu	Ser	Ser	Ser	Ser 80	Ile	Leu	Ser	Trp	Lys 85	Phe	Cys	His	Val	Leu 90
His	Lys	Val	Leu	Arg 95	Asp	Gly	His	Pro	Asn 100	Val	Leu	His	Asp	Tyr 105
Gln	Arg	Tyr	Arg	Ser 110	Asn	Ile	Arg	Glu	Ile 115	Gly	Asp	Leu	Trp	Gly 120
His	Leu	Arg	Asp	Gln 125	Tyr	Gly	His	Leu	Val 130	Asn	Ile	Tyr	Thr	Lys 135
Leu	Leu	Leu	Thr	Lys 140	Ile	Ser	Phe	His	Leu 145	Lys	His	Pro	Gln	Phe 150
Pro	Ala	Gly	Leu	Glu 155	Val	Thr	Asp	Glu	Val 160	Leu	Glu	Lys	Ala	Ala 165
Gly	Thr	Asp	Val	Asn 170	Asn	Ile	Phe	Gln	Leu 175	Thr	Val	Glu	Met	Phe 180
Asp	Tyr	Met	Asp	Cys	Glu	Leu	Lys	Leu	Ser	Glu	Ser	Val	Phe	Arg

•	WO 99	/60986											PCT/U	J S99/1
				185					190					195
Gln	Leu	Asn	Thr	Ala 200	Ile	Ala	Val	Ser	Gln 205	Met	Ser	Ser	Gly	Gln 210
Cys	Arg	Leu	Ala	Pro 215	Leu	Ile	Gln	Val	Ile 220	Gln	Asp	Cys	Ser	His 225
Leu	Tyr	His	Tyr	Thr 230	Val	Lys	Leu	Met	Phe 235	Lys	Leu	His	Ser	Cys 240
Leu	Pro	Ala	Asp	Thr 245	Leu	Gln	Gly	His	Arg 250	Asp	Arg	Phe	His	Glu 255
Gln	Phe	His	Ser	Leu 260	Lys	Asn	Phe	Phe	Arg 265	Arg	Ala	Ser	Asp	Met 270
Leu	Tyr	Phe	Lys	Arg 275	Leu	Ile	Gln	Ile	Pro 288	Arg	Leu	Pro	Glu	Gly 285
Pro	Pro	Asn	Phe	Leu 290	Arg	Ala	Ser	Ala	Leu 295	Ala	Glu	His	Ile	Lys 300
Pro	Val	Val	Val	Ile 305	Pro	Glu	Glu	Ala	Pro 310	Glu	Glu	Glu	Glu	Pro 315
Glu	Asn	Leu	Ile	Glu 320	Ile	Ser	Ser	Ala	Pro 325	Pro	Ala	Gly	Glu	Pro 330
Val	Val	Val	Ala	Asp 335	Leu	Phe	Asp	Gln	Thr 340	Phe	Gly	Pro	Pro	Asn 345
Gly	Ser	Met	Lys	Asp 350	Asp	Arg	Asp	Leu	Gln 355	Ile	Glu	Asn	Leu	Lys 360
Arg	Glu	Val	Glu	Thr 365	Leu	Arg	Ala	Glu	Leu 370	Glu	Lys	Ile	Lys	Met 375
Glu	Ala	Gln	Arg	Туr 380	Ile	Ser	Gln	Leu	Lys 385	Gly	Gln	Val	Asn	Gly 390
Leu	Glu	Ala	Glu	Leu 395	Glu	Glu	Gln	Arg	Lys 400	Gln	Lys	Gln	Lys	Ala 405
Leu	Val	Asp	Asn	Glu 410	Gln	Leu	Arg	His	Glu 415	Leu	Ala	Gln	Leu	Lys 420
	Leu	Gln	Leu	Glu 425	Gly	Ala	Arg	Asn	Gln 430	Gly	Leu	Arg	Glu	Glu 435
Ala	Glu	Arg	Lys	Ala	Ser	Ala	Thr	Glu	Ala	Arg	Tyr	Ser	Lys	Leu

v	VO 99/	60986											PCT/U	S99/11743
•	VO 331	00700		440					445					450
Lys	Glu	Lys	His	Ser 455	Glu	Leu	Ile	Asn	Thr 460	His	Ala	Glu	Leu	Leu 465
Arg	Lys	Asn	Ala	Asp 470	Thr	Ala	Lys	Gln	Leu 475	Thr	Val	Thr	Gln	Gln 480
Ser	Gln	Glu	Glu	Val 485	Ala	Arg	Val	Lys	Glu 490	Gln	Leu	Ala	Phe	Gln 495
Met	Glu	Gln	Ala	Lys 500	Arg	Glu	Ser	Glu	Met 505	Lys	Met	Glu	Glu	Gln 510
Ser	Asp	Gln	Leu	Glu 515	Lys	Leu	Lys	Arg	Glu 520	Leu	Ala	Ala	Arg	Ala 525
Gly	Glu	Leu	Ala	Arg 530	Ala	Gln	Glu	Ala	Leu 535	Ser	Arg	Thr	Glu	Gln 540
Ser	Gly	Ser	Glu	Leu 545	Ser	Ser	Arg	Leu	Asp 550	Thr	Leu	Asn	Ala	Glu 555
Lys	Glu	Ala	Leu	Ser 560	Gly	Val	Val	Arg	Gln 565	Arg	Glu	Ala	Glu	Leu 570
Leu	Ala	Ala	Gln	Ser 575	Leu	Val	Arg	Glu	Lys 588	Glu	Glu	Ala	Leu	Ser 585
Gln	Glu	Gln	Gln	Arg 590	Ser	Ser	Gln	Glu	Lys 595	Gly	Glu	Leu	Arg	Gly 600
Gln	Leu	Ala	Glu	Lys 605	Glu	Ser	Gln	Glu	Gln 610	Gly	Leu	Arg	Gln	Lys 615
Leu	Leu	Asp	Glu	Gln 620	Leu	Ala	Val	Leu	Arg 625	Ser	Ala	Ala	Ala	Glu 630
Ala	Glu	Alã	ıle	Leu 635	Gln	Asp	Ala	Val	Ser 640		Leu	Asp	Asp	Pro 645
Leu	His	Leu	a Arg	Cys 650		Ser	Ser	Pro	Asp 655		Leu	Val	Ser	Arg 660
Ala	Gln	ı Alá	a Ala	Leu 665		Ser	Val	Ser	Gly 670		Glu	Gln	Gly	His 675
Thr	Glr	туз	Leu	Ala 680		: Ser	Glu	ı Asp	Ala 685		· Ala	Leu	ı Val	Ala 690
Ala	a Lei	ı Th	r Arg	g Phe	e Ser	His	s Lev	ı Ala	a Ala	a Asp	Thr	Ile	e Val	Asn

	WO 99/60986												PCT/US99/11743		
				695					700					705	
Gly	Ala	Ala	Thr	Ser 710	His	Leu	Ala	Pro	Thr 715	Asp	Pro	Ala	Asp	Arg 720	
Leu	Met	Asp	Thr	Cys 725	Arg	Glu	Cys	Gly	Ala 730	Arg	Ala	Leu	Glu	Leu 735	
Val	Gly	Gln	Leu	Gln 740	Asp	Gln	Thr	Val	Leu 745	Arg	Arg	Ala	Gln	Pro 750	
Ser	Leu	Met	Arg	Ala 755	Pro	Leu	Gln	Gly	Ile 760	Leu	Gln	Leu	Gly	Gln 765	
Asp	Leu	Lys	Pro	Lys 770	Ser	Leu	Asp	Val	Arg 775	Gln	Glu	Glu	Leu	Gly 780	
Ala	Met	Val	Asp	Lys 785	Glu	Met	Ala	Ala	Thr 790	Ser	Ala	Ala	Ile	Glu 795	
Asp	Ala	Val	Arg	Arg 800	Ile	Glu	Asp	Met	Met 805	Ser	Gln	Ala	Arg	His 810	
Glu	Ser	Ser	Gly	Val 815	Lys	Leu	Glu	Val	Asn 820	Glu	Arg	Ile	Leu	Asn 825	
Ser	Cys	Thr	Asp	Leu 830	Met	Lys	Ala	Ile	Arg 835	Leu	Leu	Val	Met	Thr 840	
Ser	Thr	Ser	Leu	Gln 845	Lys	Glu	Ile	Val	Glu 850	Ser	Gly	Arg	Gly	Ala 855	
Ala	Thr	Gln	Gln	Glu 860	Phe	Tyr	Ala	Lys	Asn 865	Ser	Arg	Trp	Thr	Glu 870	
Gly	Leu	Ile	Ser	Ala 875	Ser	Lys	Ala	Val	Gly 888	Trp	Gly	Ala	Thr	Gln 885	
Leu	Val	Glu	Ser	Ala 890	Asp	Lys	Val	Val	Leu 895	His	Met	Gly	Lys	Tyr 900	
Glu	Glu	Leu	Ile	Val 905	Cys	Ser	His	Glu	Ile 910	Ala	Ala	Ser	Thr	Ala 915	
Gln	Leu	Val	Ala	Ala 920	Ser	Lys	Val	Lys	Ala 925	Asn	Lys	Asn	Ser	Pro 930	
His	Leu	Ser	Arg	Leu 935	Gln	Glu	Cys	Ser	Arg 940	Thr	Val	Asn	Glu	Arg 945	
Ala	Ala	Asn	Val	Val	Ala	Ser	Thr	Lys	Ser	Gly	Gln	Glu	Gln	Ile	

WO 99/60986			PCT/US99/11743
-------------	--	--	----------------

'	NO 99/	60986		_									1 01/0	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
				950					955					960
Glu	Asp	Arg	Asp	Thr 965	Met	Asp	Phe	Ser	Gly 970	Leu	Ser	Leu	Ile	Lys 975
Leu	Lys	Lys	Gln	Glu 980	Met	Glu	Thr	Gln	Val 985	Arg	Val	Leu	Glu	Leu 990
Glu	Lys	Thr	Leu	Glu 995	Ala	Glu	Arg	Val	Arg L100	Leu	Gly	Glu	Leu	Arg 1105
Lys	Gln	His		Val 1110	Leu	Ala	Gly	Gly	Met 1115	Gly	Thr	Pro	Ser	Glu 1120
Glu	Glu	Pro		Arg 1125	Pro	Ser	Pro	Ala	Pro 1130	Arg	Ser	Gly	Ala	Thr 1135
Lys	Lys	Pro		Leu 1140	Ala	Gln	Lys		Ser 1145	Ile	Ala	Pro	Arg	Thr 1150
Asp	Asn	Gln		Asp 1155	Lys	Lys	Asp		Val 1160		Pro	Ala	Gln	Leu 1165

Val Asn Tyr

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 12:
- GAAGATACCC CACCAAAC 18
- (2) INFORMATION FOR SEQ ID NO:13:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: other DNA
- (iii) HYPOTHETICAL: no



- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GCTTGACAGT GTAGTCATAA AGGTGGCTGC AGTCC 35

- (2) INFORMATION FOR SEQ ID NO:14:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: other DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 14: GGACATGTCC AGGGAGTTGA ATAC 24
- (2) INFORMATION FOR SEQ ID NO:15:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 41
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: yes
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CUACUACUAC UACUAGGCCA CGCGTCGACT AGTACGGGII GGGIIGGGII G 41

- (2) INFORMATION FOR SEQ ID NO:16:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 516
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human

/11743

WO 99/60986	PCT	/US99/1
(x) FEATURE: exon 1 of HIP1		
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 16: TCTGTGGAAG GTTTGGAGGG GAGAGAGGGG CAGCTGGATG CTCTTGGGCC	∆CGGTCGCCC	60
TCTGTGGAAG GTTTGGAGGG GAGAGAGGGG CAGCTGGATG CTCTTGGGCC CTGATCTCTG CGCCTCTTCC TCCTGCTCCG GGAGAAATAA TGTTTCCCTG		120
GCATCTCTTT GTGCGGGCTT TAATTGCCAT GTTGTTGTGC CAAGGGAGTG		180
GGACCAGCAG CTGGGCACAG CCAATGCCAG GCAGTGGTGC CCACTCCCTC		240
GCCAGCTGGC TCCTGGGAGC GCTGCCCACC TCTGCCCCCA GCTGGGCGCC		300
CGACCACCCG TGGGGCTGGG GGAGGTTGGC TGGAGGAGGA GAAAGGGGCG		360
GGGTCTCAGC CACTCTCAGA GGCTTATTCA TCTCATCCTC CTTTCCCTCC		420
TTTTCAGACT GTCAGCATCA ATAAGGCCAT TAATACGCAG GAAGTGGCTG		480
ACACGCCAGA AATATCCTTT GGATGTTGCT TGGAAG		516
(2) INFORMATION FOR SEQ ID NO:17: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 193 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 2 of HIP1 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17: TGTTTTCCAT AACCCCCCT CACCGTGCAT ACTGGGCACC CACCATGAGA GACCTTCTGG TCTGTTGTCA ACCGCCTGCC TCTGTCTAGC AACCCAGTGGG GTTCTGCCAT GTGTTCCACA AACTCCTCCG AGATGGACAC CCGAACGTGA	TCTGCTGGAA	60 120 180
(2) INFORMATION FOR SEQ ID NO:18: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 104 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 3 of HIP1 (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 18: GTGTTCTTTT GCCCCTGCAG GTCCTGAAGG ACTCTCTGAG ATACAGAAAC		193
ACATGAGCAG GATGTGGGTG AGTTTGGAGA TGTACTCAGG AGCC		104

- (2) INFORMATION FOR SEQ ID NO:20:
- (i) SEQUENCE CHARACTERISTICS:



(A) LENGTH: 327

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human

(x) FEATURE: exon 4 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AATTCCTGGC TGCAGATCTC TTGACTGTTA TGTTCTTGTT GTTGACTCTG TTTCCCCTCC 60
TCTTCCTAAA AGGGCCACCT GAGCGAGGG TATGGCCAGC TGTGCAGCAT CTACCTGAAA 120
CTGCTAAGAA CCAAGATGGA GTACCACACC AAAGTGAGTC TCTGCGGACA GTTCTGCCGC 180
CACCGCCGCC TCCCCTGCTC CATCCCTTCA GCCCCTCCCT GGGCTCATTT GTCAGCTCTT 240
TCAGGTAATA GACAGCCCAG GCCTCCGG AAGTGTGCAC ATCATGTACC CAAGCTGTGA 300
GAGAGGAAAG CCACCGCCAG GCCCACG 327

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 331(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human

(x) FEATURE: exon 5 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGGCTCAAGC AATCCTCCCA CCTCGGCCTC CCAAGTAGCT GGGACCACAG GCGTGTGCCA
CCACGCCCGG CTGAGAGAGG GCTCTTCATG TCTTCTGCCC TGACTCCCTT CCTCTGCCTC
CCTTCCAGAA TCCCAGGTTC CCAGGCAACC TGCAGATGAG TGACCGCCAG CTGGACGAGG
CTGGAGAAAG TGACGTGAAC AACTTGTAAG TGGCTCCTGC CCTGAGCCCA GGGAGGGAGA
AAGCTTTTGT GAATGCTGAC ACTTCTCATA AGGGTCATGG AGGGCCTGAT GGGGGGAGGC
CGTGGCTGGG ATGGGGACCA AAGCCCCTGG G
331

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 470
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no

WO 99/60986

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FEATURE: exon 6 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 22:

` '						
ACTGTCGCTG	TCACTGTTGA	CTTCACCAGG	CTGCATGGCC	${\tt ATAATACCCA}$	CAAGGCTAAG	60
ACTTGGAGCT	GGAGTTGTGT	GTGTGTTTGC	GCATGCACAT	GAGCATTGGA	GACTGGAGTA	120
GCGTAGAGCG	TGGGGGAGGG	GACAGGTAAC	AGACCGGCCT	CAGGCTGTGG	AGTGTAAGCT	180
CTCTTTCCTC	TTGGGTCCAG	TTTCCAGTTA	ACAGTGGAGA	TGTTTGACTA	CCTGGAGTGT	240
				GTCTAACCCA		300
				TGCAGCACAA		360
				TCGTGCAGGG		420
				CGGGTGCAGT		470

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 565

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FEATURE: exon 7 of HIP1

(xi)SEOUENCE DESCRIPTION: SEQ ID NO: 23:

TCTTCACCTG	TTTAATGGGG	ATACGTTTAC	CTATCTCATG	GGAGTGTTGT	GAAGGTTAAA	60
TCAATTAGAT	GAGGTAAAGC	ACGCACAGAA	TCGGTCCTTG	${\tt GTGTATGTTG}$	GACCCCTGCC	120
TOTALLICATION	TGAAGAGGCT	GCCTGTAATC	CCCTGGCTCT	ACCACCTTTC	TCCCTCACTT	180
TCTGCCCTC	CHATTCAACT	CCCTGGACAT	GTCCCGCTCT	GTGTCCGTGA	CGGCAGCAGG	240
						300
				TGCAGCCACC		
CACTGTCAAG	CTTCTCTTCA	AACTCCACTC	CTGTGAGTAC	CGCGGGCCAG	ATCTTCTTAC	360
				TCCCCAGAGA		420
				CCCGGAGCTT		480
						540
TGAGGATAAA	AGAGCAGGGC	CCAGGCTTTG	GTGACCCCAG	TAAAGCCCCT	GGCTTGCCAC	540
						565
TCTTGCGTCC	AGTGTTACAG	GAICI				

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 233

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FEATURE: exon 8 of HIP1

PCT/US99/11743

WO 99/60986	

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GGGACAGCTC	TAGGCCAGTC	GTGGCCCCTG	GCAGTGCTGG	CCACATGCCC	CAGGGTAGCT	60
GGGCCCCTCC	CCCTCGAGAG	CCCCGCTGTG	GCTTCCCTGC	CCTCTGGTCC	CCCTCCCCTC	120
TCACACTCTT	TCCAATTTCT	TCCAGGCCTC	CCAGCTGACA	CCCTGCAAGG	CCACCGGGAC	180
CGCTTCATGG	AGCAGTTTAC	AAAGTAAGTG	GTTCAAGTAA	CAGGAATGGA	GGT	233

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 578

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FEATURE: exons 9 and 10 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TGAATCCCAG	CACCATGGAG	${\tt TTTATCTCCT}$	TGACAGCCTG	TGCCTTTGGG	CTGGGGAGGG	60
GGCAGGAAAG	CCAGGTGGCT	GCTCTGTCCC	${\tt CTACATGGGG}$	CTGATGAAGA	CACCCAGCAC	120
CCCTCAGGTC	CTTCTCCACC	CCTAGGTTGA	${\tt AAGATCTGTT}$	CTACCGCTCC	AGCAACCTGC	180
AGTACTTCAA	GCGGCTCATT	CAGATCCCCC	AGCTGCCTGA	GGTAAGCATG	CCCAACCACA	240
CACCCTCGGC	ACTGCAGAGG	CCCCAGGTAC	TCTCTTAAGG	GCCGGCGGG	CCTGGCAAGC	300
AAGCACTATT	${\tt TGAGGATGTG}$	TCTCCGTCTT	CAGAACCCAC	CCAACTTCCT	GCGAGCCTCA	360
GCCCTGTCAG	AACATATCAG	CCCTGTGGTG	GTGATCCCTG	CAGAGGCCTC	ATCCCCGAC	420
AGCGAGCCAG	TCCTAGAGAA	GGATGACCTC	ATGGACATGG	ATGCCTCTCA	GCAGGTGAGG	480
ACCACTTGGG	AGAGAAACTT	${\tt GGCCTTTCCT}$	${\tt CTCACCTGCA}$	AGTACAGGGG	AGAGGCTGGG	540
${\tt GGAGACCCTG}$	GCCAAAGCCC	${\tt ATTGACTCTA}$	ACCAGGTT			578

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 390

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no(iv) ANTI-SENSE: no(vi) ORIGINAL SOURCE:(A) ORGANISM: human(x) FEATURE: exon 11 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 26:

TTTAAAAAAA	AAAAAATTAA	ACAGGTCTGA	${\tt ACCGTTTAAT}$	TCGAGAAAGG	GGGCATTCTC	60
CCATATCACT	CAACTGACCC	ACACACAGAA	${\tt TTCTCTGGCT}$	CTCTGACTTA	TTCTCACTCC	120
TTTTTGGTCA	ACCACAGAAT	${\tt TTATTTGACA}$	ACAAGTTTGA	TGACATCTTT	GGCAGTTCAT	180
TCAGCAGTGA	$\mathtt{TCCCTTCAAT}$	TTCAACAGTC	AAAATGGTGT	GAACAAGGAT	GAGAAGTGAG	240
TCCAAGCTGG	${\tt GTTCAAGCAG}$	ATGGTTCAGG	AGCTAAGTTA	AGCCATGGTC	TGCCTCAAAA	300
CACTAACCAA	AGAGGAATTC	TTAATGATAC	TGGGGCTTCT	TAGATACAGA	ACATCTTGAA	360
GGGTTGGGGG	CAATGGCTTA	TGCCTGTAAT				390

PCT/US99/11743 WO 99/60986

WO 35/00500	
(2) INFORMATION FOR SEQ ID NO:27: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 547 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 12 of HIP1 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27: AAAATCAATA ACCATGGATT TATGAGTATT AGATTAGTAT TCCTCCCTG AGCTCACAAT TAAAACAGA GGGATCAAAT CACTATGAAA CCACTATTAATT TCCTCCCCTG TCCCCAGGGA CCACTTAATT GAGCGACTAT ACACTAGAAA GCAAACTCAT TCCCCTTCCC TCCCCAGGGA CCACTTAATT GAGCGACTAT ACACTAGCAA GGCACAGC TTCACCAAAA CACCTAGCCA GGCACAGCC TAATCCTAGC ACTTTGGGAG GCCAAGGCAG GAGGATTACC TGAGGTCGGG AGTTCGACCTG TAATCCTAGC ACTTTGGGAG CCCACGCTCT TCCAATAAAA ATGCAATAAT TAGCCGGGTG TGTTGGCAG CACCTGTAAT CCCCAGCTACT CGGGAAGCTG AGGCATGAGA ATTGCTTGAA CTTGGGA	60 120 180 240 300 360 420 480 540 547
(2) INFORMATION FOR SEQ ID NO:28: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 436 (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
(D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA	
(iii) HYPOTHETICAL: no	
(iv) ANTI-SENSE: no	
(vi) ORIGINAL SOURCE: (A) ORGANISM: human	
(x) FEATURE: exon 13 of HIP1	
(xi)SEOUENCE DESCRIPTION: SEQ ID NO: 28:	
CCCCCAGCCA CTCTAAAGAG GACCACAATT CCCCGGCCAT CATCCCCTGT TATTGTTGTT	60
GATTGAGGGG CTCCTAATGA CCAGATGGTC CAACCCTCCT GGGACGTGGA GAGTTGACTT	120 180
AGGGGAATCA GGTATTTACT TGGAAGCATG GTAGGACCCG CTTCTCCGGC CCATGCCCGT	240
GACCCGTGGC AGTGGGCGGT TGGCCTCATG ACCGGAGTCC CCCCACAGAG CCAGCGGGTT	300
GTGCTGCAGC TGAAGGGCCA CGTCAGCGAG CTGGAAGCAG ATCTGGCCGA GCAGCAGCAC CTGCGGCAGC AGGCGGCCGA CGACTGTGAA TTCCTGCGGG CAGAACTGGA CGAGCTCAGG	360
CTGCGGCAGC AGGCGGCCGA CGACTGTGAA TTCCTGCGGG CAGAACTGGA CGAGCTCAGG AGGCAGCGGG AGGACACCGA GAAGGCTCAG CGGAGCCTGT CTGAGATAGA AAGTGAGCGG	420
AGGCAGCGGG AGGACACCGA GAAGGCICAG CGGAGCCIGI CIGAGAIAGA IMOIGH	436

- (2) INFORMATION FOR SEQ ID NO:29:
- (i) SEQUENCE CHARACTERISTICS:

TGGGTGGGG CGGGGG



(A) LENGTH: 469

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FEATURE: exon 14 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CACTTGAGCC	CAAGGAGGTC	AAGGCTGCAG	TODDOOD	TO TO THE TOTAL CONTRACTOR OF	CCACCCCACC	60
						60
CTGGGTGACA	GAGCAAGACT	GTCTCAAAAC	AAAACAAGGA	GGACCTTCTA	GGGACCCTGG	120
CTCATTGCAA	GGAAGGCAAG	GGTCCCTGCT	AGGTTAGACT	CCTCACCTTG	GTCCTTTACA	180
ATACAGGGAA	AGCTCAAGCC	AATGAACAGC	GATATAGCAA	GCTAAAGGAG	AAGTACAGCG	240
AGCTGGTTCA	GAACCACGCT	GACCTGCTGC	GGAAGGTAAG	ACCCTCAGCC	CCTGTCACCA	300
TCCTGCAGGC	CCTGCACCTC	TAGGGAGAGA	GCGGCTCAGG	CCTGTGGCTT	CCCCGGGGCC	360
AGCAACCCCT	ACATTGATCT	CTAAGGCATT	GCCGTCATCT	CGGGAACCAC	ACCTTTTCAG	420
GCTTCCTTGC	CTCTGTGTCT	TGGGCTGTGT	CCTGGGTGCC	AATCCCATG		169

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 359

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FEATURE: exon 15 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GGGTAGGAAA	${\tt GTGATTCCTG}$	${\tt TGTCTGACTC}$	TAGGGCACGC	ACAGCCTGAG	TATGATTGTC	60
CTAGAAGGAG	GATGTCCTCT	AAGCCTGGGA	${\tt TCTCCTGGTT}$	CAAGACACTG	TTCTTCTTTT	120
GCAGAATGCA	GAGGTGACCA	AACAGGTGTC	CATGGCCAGA	CAAGCCCAGG	TAGATTTGGA	180
ACGAGAGAAA	AAAGAGCTGG	${\tt AGGATTCGTT}$	GGAGCGCATC	AGTGACCAGG	GCCAGCGGAA	240
GGTGAGTGGG	ACGAGGAGCA	CTCGGGAAAT	GAGGGAGGG	GCTGTTGAGT	TGGTGGCGGG	300
CCCTTTCTCC	ССФФСФССФС	CATGGGCAGT	тететесете	COMPCCCATC	ACACACCAC	350

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 209

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

209

(iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 16 of HIP1	
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 31:	
GTTGATCGCT TGGGACGTTT TTACATTTTT ATATTCTTTG TCACTGTCAC CCAGATCAGA	60
GTCCCTCTGT TTTTCTTCTC TTTCAGACTC AAGAACAGCT GGAAGTTCTA GAGAGCTTGA	120
AGCAGGAACT TGCCACAAGC CAACGGGAGC TTCAGGTTCT GCAAGGCAGC CTGGAAACTT	180

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

CTGCCCAGGT AAATACCTCC TTTTTTTT

- (A) LENGTH: 485
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 17 of HIP1
- (xi)SEOUENCE DESCRIPTION: SEQ ID NO: 32:

(XI)SEQUEIN	CE DESCRI	LION. SEQ	10 10. 32.			
CCCCCACTGC	AATCAGTGTG	TCCCCGGGAG	GGAATCAGAG	${\tt TGGCAGGTTA}$	AAGAGCCATC	60
* CCCCCCCCCC	TCCTTCC A AC	CCGGTGGTGG	GTTGGACCTC	${\tt TGGGAAGTAG}$	GGACTGTTTA	120
ACCITCCCAG	TCCTTGCAAC	mmmccmmcmc	CTCACCTTTC	CAGTCAGAAG	CAAACTGGGC	180
ACTCAACCAG	CGTCTCCCTC	TTTCCTTGIG	GICACCITIG	cman amagaa	CACCECATAC	240
AGCCGAGTTC	GCCGAGCTAG	AGAAGGAGCG	GGACAGCCTG	GTGAGTGGCG	CAGCICATAG	
GGAGGAGGAA	TTATCTGCTC	TTCGGAAAGA	ACTGCAGGAC	ACTCAGCTCA	AACTGGCCAG	300
CACACACCCT	CACGGACATG	GACACGAGCG	AGCACCTGTG	AATTCCCACC	GAGGGCCTCT	360
CACAGAGGGI	OTHORDOOM	ACCACCCCC	CCCTCCTCAC	Δ Δ G G G G T T T G	GGGCCTTGGC	420
GCGCATGCAC	GGAGGCTGGG	AGGACCCCGG	GGC1GC1GAG	OCTOOD TO	CTCCACACTC	480
CTGATTGTGC	AGACATTCTG	TAGGTGTAAT	GCCAGCAGGC	CCTGCATTGC	CTGCAGAGTC	
CATGA						485
O						

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 468
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 18 of HIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 33:

TTACTGGCTT GGCCATGACT TGAGCTAAGA TGCTAAGAGC CCCAGCCAGG
TCATCCTGCT CAGGTTCATT ATGGAGTCTA GGGCAGACTC TCACCTCCCT GGACCATTTT
120

WO 99/60986		PCT	/US99/11743
TAGAATCTAT GTGCCAGCTT GCCAAAGACC AACGAAAAAT			180
AGGCTGCGGA GCAGGTGATA CAAGACGCC TGAACCAGCT			240
GCTGCGCTGG GTCTGCAGGT ACACTTGCAA TTGCCCAGCT AGCCTGAGAC TCTGTTGATG TTGAATCTCA TGTGAGACTT			300
AGCAGCATGT CAGCATTACC TTAGGGGCGC CCAGGCCCCA			360 420
GAAACTCTGT GCATTAGTGC CTATACACTA GTATTTTAGT		01111011010	468
(2) INFORMATION FOR SEQ ID NO:34:			
(i) SEQUENCE CHARACTERISTICS:			
(A) LENGTH: 393			
(B) TYPE: nucleic acid			
(C) STRANDEDNESS: double			
(D) TOPOLOGY: linear			
(ii)MOLECULE TYPE: genomic DNA			
(iii) HYPOTHETICAL: no			
(iv) ANTI-SENSE: no			
(vi) ORIGINAL SOURCE:			
(A) ORGANISM: human			
(x) FEATURE: exon 19 of HIP1			
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 34: CACTAGTAAG CTCCTCCATT CAGTGCTTAA TTAACGAGGA	man naganga	#3.#G3.G3.7.0m	60
TGCTCTGACC TTGCCCTGTG TTCCCTCTCA CAGATCACCT			60 120
TTTCCAGCTG CATCGAGCAA CTGGAGAAAA GCTGGAGCCA			180
GTAAGAATGG CCAAGGACAG TCTCTGTCGG CTAGTGATGG			240
CCTGAATGCG GGGATAGTGA CAGGTCCCTC TGCATCAAGA	AAGGCATGTA	GGCAACTCAT	300
ACAAGAAAGG CATGTAGGCA ACTCATAAAA CGGGAGGAGA	GGGTATGAAA	GTGTCACCAT	360
CAACCAGACC TGAGAAACTT CTCTTTCCAA TCC			393
(2) INFORMATION FOR SEQ ID NO:35:			
(i) SEQUENCE CHARACTERISTICS:			
(A) LENGTH: 421			
(B) TYPE: nucleic acid			
(C) STRANDEDNESS: double			
(D) TOPOLOGY: linear			
(ii)MOLECULE TYPE: genomic DNA			
(iii) HYPOTHETICAL: no			
(iv) ANTI-SENSE: no			
(vi) ORIGINAL SOURCE:			
(A) ORGANISM: human			
(x) FEATURE: exon 20 of HIP1			
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 35:	macaama ama	AMCCCCCA CA C	60
GGCCTGCCCA GAAGGTAAGA ATGGCCAAGG ACAGTCTCTG AGGGTTCAGA AGCACCTGAA TGCGGGGATA GTGACAGGTC			60 120
TGTAGGCAAC TCATACAAGA AAGGCATGTA GGCAACTCAT			180
GAAAGTGTCA CCATCAACCA GACCTGAGAA ACTTCTCTTT			240
TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC			300
MCCCACCACC MCCCMCACAC CCCCACCMCA CCCMCCCCCA	mama a ama am	CCCCCAMCAC	3.60

TGCCACCACC TGCCTCAGAG CCCCACCTGA GCCTGCCGAC TGTGAGTACT GGGGCATGAG

GGGCTGTTCA TGGACCAGGG GAGCAGGGGG CCTTTAAAAG TCTCTGTTGG GCCGGGCGCA

G

360

420

421

PCT/US99/11743

WO 99/60986

- (2) INFORMATION FOR SEQ ID NO:36:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 498
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 21 of HIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 36:

ACCCCCACCC	ACCACAATCG	CTTGAACTCA	GGAGGCGGAG	${\tt TTTGCAGTGA}$	GCCGAGATGG	60
CCCCACTCCA	CTCCAGCCTG	GGCAACAAGA	GCGAGACTCC	ATCTCAAAAA	AAAAGTGTCT	120
				GTATGGCAGG		180
				TGCCGACAGC		240
CCIACCTCCT	GAGCAAGATC	AAGGCCATCG	GCGAGGTACT	TGGAGTAGTA	TCATTGAGGA	300
CCATTGTTAT	TCTTCTGGGT	GTGCGTGCTG	GTGAATGGCC	AGGGAATCGG	TGATGTTCTG	360
ACCTACTTCT	TTCTGCACTT	AGAACTTGAT	TCTAGAAAGA	GATTGTTAAA	ATTGGAAAAT	420
CTGGCCGGGT	GCAGTGATTT	ATGCGTGTAA	TCCCAGCACT	TTGGGAGGCC	GAGTCAGGAG	480
GATCACTTGA						498
Gritchellon						

- (2) INFORMATION FOR SEQ ID NO:37:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 427
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 22 of HIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 37:

()		GTTTGCTGGG	mccccmccmc	$CCTTCCC\DeltaTC$	TTGTAAGGGT	60
CCCTGTGGCT	TGCAGAAGGT	GT-T-TGC-TGGG	1666616616	CCIIGCCAIC	11011210001	400
TACAGATGGC	AGAGGAGAAG	AGACAGGAGG	CCCCAAGGTC	AGTTCAGCCT	TTGTGATGTG	120
Inchionicoc			CAMCAACCAC	CACCACCTCC	CCCACCTCGT	180
TTCACAGGAG	CTCCTGCCCA	GGGGACTGGA	CATCAAGCAG	GAGGAGCIGG	00011001001	240
CCACAACCAC	ATRICCICCICCA	CTTCAGCTGC	TATTGAAACT	GCCACGGCCA	GAATAGAGGT	240
GGACAAGGAG	MIGGCGGCCII	CITCHOOLG	maaammmaa A	aamaccammc	ጥሮር እ እር እርጥር	300
AGGAGGTTCC	TGCAGGATCT	CCTGAAACGA	TGCCTTTGCA	GCTGCCCTTC	IGCAACACIG	• • • •
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	A MOMO A C A CM	CGTTCATTAA	CCCCATGCCA	ACCCCCTAAG	ACAGAAACCA	360
CTCATTAAAC	ATGICACAGI	CGIICAIIAA	000011100011		armar ammar	420
CAATTTCCCA	GGCACAGTGG	CTCATGCCTG	TAACCCCAGC	ACCTTGGGAG	GATCACTIGA	420
GAATITOCCII	000					427
GTCCAGG						

- (2) INFORMATION FOR SEQ ID NO:38:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 367

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FEATURE: exon 23 of HIP1

# (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 38:

CCCCTGAAT	AGGTTAGAGT	CTGGATTCTT	TTCTGACTCT	CTCAAGAATG	TGGGCAGGGA	60
CTTGGGGACT	${\tt TCCAGATTCA}$	GGTTTCCCAG	CTACCACACG	ATGTTGGACT	GAAAGTATAG	120
TAAGACATTA	${\tt GTGGATCCTT}$	AATATTCAAG	${\tt GCACATTTAG}$	AAACCATGCT	TCTTTTTCAC	180
AGGAGATGCT	CAGCAAATCC	CGAGCAGGAG	ACACAGGAGT	CAAATTGGAG	GTGAATGAAA	240
GGTCGGTCTG	AGCGGCATGG	TGGGACCTAG	GGGAGCAGGA	TCTGTCTTCC	TGACATTGGT	300
CTATACTTTG	${\tt CATACTTATT}$	AGGGAATTAG	AGGAGAGCAG	TAGCAGCCAC	GGGGAAGGGC	360
TGAGTTG						367

# (2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 502
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 24 of HIP1

# (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 39:

CCCCGCAGAA	TGTTCCAGCA	ACCTCAGCAC	CCTTCTTACC	TCCCTTTCCC	ATTCCAAGCT	60
TGCCTTTGGC	TAGGAGTGGG	GAAGAGAACC	${\tt GTCGTGTTCA}$	TTGATCTTGG	ATCTTGATCT	120
CAGTGTATCC	TCGACTTGTT	TGTTTGGCAG	${\tt GATCCTTGGT}$	TGCTGTACCA	GCCTCATGCA	180
AGCTATTCAG	GTGCTCATCG	TGGCCTCTAA	GGACCTCCAG	AGAGAGATTG	TGGAGAGCGG	240
CAGGGTGAGC	GTGGGTGTGG	GCCCTGGGCA	GGAAGAGGAG	GCATCGGTGA	CAGACTCCCG	300
CTCCAACGGA	CTCTGTGATG	CTGCCGTCTT	ACTCTGTGTG	TCCACCTGAG	TACAGAGCAG	360
CCACTCCTGT	AGATATCAGC	AGAGGCCCTG	GGGAGAAGTC	AGAGCTCCAG	GACCTCCCCA	420
GAGGGTGGCC	AGGCATGTGT	CCCAACTCCA	GCTCCCTTCG	CACAGGCAGA	CATTGTTGGA	480
ACTTGCTGTG	GGAGCCCTTT	TT				502

# (2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 437
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no

437

(vi) ORIGINA	L SOURCE	:				
(A) ORGANI	SM: human					
(x) FEATURE	E: exon 25 of	HIP1				
(xi)SEQUENC	CE DESCRI	PTION: SEQ	ID NO: 40:			
TTTTGGTCTC T	GAATCTTCT	TCTTTTTTGT	AAAATGGGAA	${\tt TACTAATGCT}$	TATGTCTCAG	60
AGTTACTATG A	AGGATGATTT	GGGATAATAT	ATGTATAAAA	${\tt GCACCTGCCA}$	TATAGTACAT	120
GCTCAATAAA A	AGGTGGCTAT	TACTATTTTT	${\tt TATTTCCCTA}$	GGGTACAGCA	TCCCCTAAAG	180
AGTTTTATGC C	CAAGAACTCT	CGATGGACAG	AAGGACTTAT	CTCAGCCTCC	AAGGCTGTGG	240
GCTGGGGAGC C	CACTGTCATG	GTGTAAGTAT	CTATTGGTAC	${\tt CAAGGGTCCT}$	CCCATGACCC	300
CTCTTCCATT G	SATCCACTCC	AAACAATAGC	TAAGGAGGGA	AAAAAAAATC	TGTCCCTTAG	360
АААТАААСТА Т	TTGATCAGGA	AGTCAATAGG	ACCGAGTTTA	CAAGGGAGCC	TGGCTCTCCC	420

# (2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 351

AGGGGACACA GGGCAGG

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 26 of HIP1

# (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 41:

(111)022 2 02.						
GGGAGCCTGG	CTCTCCCAGG	GGACACAGGG	CAGGCAGCCT	CCCCTCCCTG	TTTAGCCAAG	60
GGCGATGGGG	TGGTCTGGAG	GTGGGATTGT	GGAGGAGTTG	CAGCTCATTT	GCCCGTAACC	120
TAGTCCCTCT	TGTCGTTTTC	CATCAGGGAT	GCAGCTGATC	TGGTGGTACA	AGGCAGAGGG	180
AAATTTGAGG	AGCTAATGGT	GTGTTCTCAT	GAAATTGCTG	CTAGCACAGC	CCAGCTTGTG	240
		TGGCTGGACC				300
		GTACTTGGCT				351

# (2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 418
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 27 of HIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 42:

CTTTTTATAT G	GATAGATATG	TCAGGAGCTG	ACTATAGTCA	GCAGATTTTG	AGAAGCTGAT	60
TGGTGATTGC C	CGTTTGGCCC	ACATATGTTT	GCTAAGAACC	ATCAGAGCAA	TTATCTGATT	120
CAGTCCTTGT T	гсстстасст	GTTGTATGAA	CCTAAATCTG	CTTTGTCCTG	GTAGGTGAAA	180

GCTGATAAGG	ACAGCCCCAA	CCTAGCCCAG	CTGCAGCAGG	CCTCTCGGGG	AGTGAACCAG	240
GCCACTGCCG	GCGTTGTGGC	CTCAACCATT	TCCGGCAAAT	CACAGATCGA	AGAGACAGGT	300
AGCCTTTCCA	AAGGGACCCT	TTTCTTACCC	ACCCTGTTGA	GCTCTTCTCT	GCATCCTTCC	360
CTGTGATCCC	AACCAAATCC	CACAGGACTG	TGTCTAAATT	CTTTCATATT	TTTCATCT	418

# (2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 279
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 28 of HIP1

# (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 43:

` '						
TTTCCACAGA	GCATTGGCAT	TGGCTGCCTC	TCAGGTGCCA	GTCAGCCAGG	GTAGAATTTG	60
ATGAGACCTT	CTTGTTTCCA	TCCTTGCAGA	CAACATGGAC	TTCTCAAGCA	TGACGCTGAC	120
ACAGATCAAA	CGCCAAGAGA	TGGATTCTCA	GGTTAGGGTG	CTAGAGCTAG	AAAATGAATT	180
GCAGAAGGAG	CGTCAAAAAC	TGGGAGAGCT	TCGGAAAAAG	CACTACGAGC	TTGCTGGTGT	240
TGCTGAGGGC	TGGGAAGAAG	GTAAGCTGAC	TCAAAGGAT			279

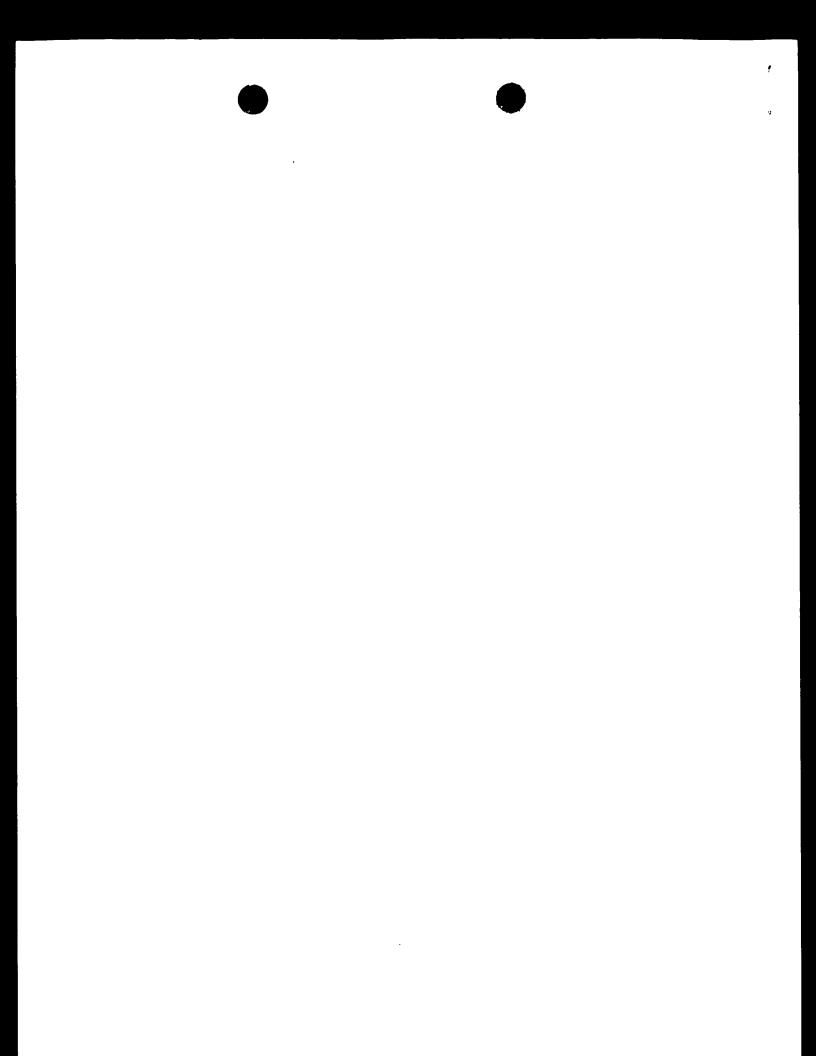
# (2) INFORMATION FOR SEQ ID NO:44:

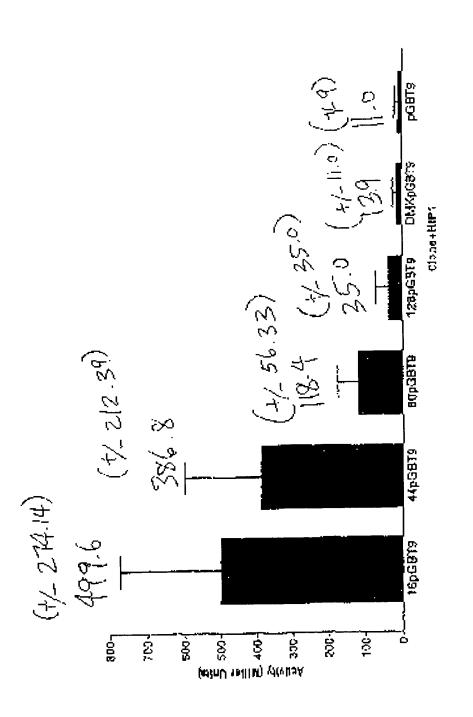
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3715
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 29 and partial cds of HIP1

# (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 44:

AACATAAATT	ATCATTGTCT	TTTAGGAACA	GAGGCATCTC	CACCTACACT	GCAAGAAGTG	60
GTAACCGAAA	AAGAATAGAG	CCAAACCAAC	ACCCCATATG	TCAGTGTAAA	TCCTTGTTAC	120
CTATCTCGTG	TGTGTTATTT	CCCCAGCCAC	AGGCCAAATC	CTTGGAGTCC	CAGGGGCAGC	180
CACACCACTG	CCATTACCCA	GTGCCGAGGA	CATGCATGAC	ACTTCCCAAA	GACTCCCTCC	240
ATAGCGACAC	CCTTTCTGTT	TGGACCCATG	GTCATCTCTG	TTCTTTTCCC	GCCTCCCTAG	300
TTAGCATCCA	GGCTGGCCAG	TGCTGCCCAT	GAGCAAGCCT	AGGTACGAAG	AGGGGTGGTG	360
GGGGGCAGGG	CCACTCAACA	GAGAGGACCA	ACATCCAGTC	${\tt CTGCTGACTA}$	TTTGACCCCC	420
ACAACAATGG	GTATCCTTAA	TAGAGGAGCT	GCTTGTTGTT	TGTTGACAGC	TTGGAAAGGG	480
AAGATCTTAT	GCCTTTTCTT	TTCTGTTTTC	TTCTCAGTCT	TTTCAGTTTC	ATCATTTGCA	540
CAAACTTGTG	AGCATCAGAG	GGCTGATGGA	TTCCAAACCA	GGACACTACC	CTGAGATCTG	600
CACAGTCAGA	AGGACGGCAG	GAGTGTCCTG	GCTGTGAATG	CCAAAGCCAT	TCTCCCCCTC	660
TTTGGGCAGT	GCCATGGATT	TCCACTGCTT	CTTATGGTGG	TTGGTTGGGT	TTTTTGGTTT	720
TGTTTTTTT	TTTTAAGTTT	CACTCACATA	GCCAACTCTC	CCAAAGGGCA	CACCCTGGG	780
GCTGAGTCTC	CAGGGCCCCC	CAACTGTGGT	AGCTCCAGCG	ATGGTGCTGC	CCAGGCCTCT	840

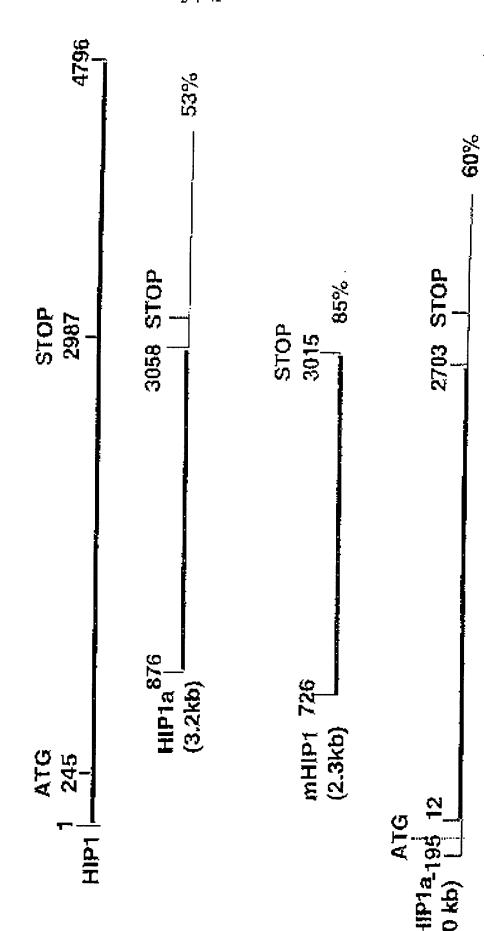
CGGTGCTCCA	TCTCCGCCTC	CACACTGACC	AAGTGCTGGC	CCACCCAGTC	CATGCTCCAG	900
GGTCAGGCGG	AGCTGCTGAG	TGACAGCTTT	CCTCAAAAAG	CAGAAGGAGA	GTGAGTGCCT	960
TTCCCTCCTA	AAGCTGAATC	CCGGCGGAAA	GCCTCTGTCC	GCCTTTACAA	GGGAGAAGAC	1020
AACAGAAAGA	GGGACAAGAG	GGTTCACACA	GCCCAGTTCC	CGTGACGAGG	CTCAAAAACT	1080
TGATCACATG	CTTGAATGGA	GCTGGTGAGA	TCAACAACAC	TACTTCCCTG	CCGGAATGAA	1140
CTGTCCGTGA	ATGGTCTCTG	TCAAGCGGGC	CGTCTCCCTT	GGCCCAGAGA	CGGAGTGTGG	1200
			CTGCCTTGGC			1260
CGTTCCACTT	TCTACGCAAT	TGACAAACCC	GGAAGATCAG	ATGCAATTGC	TCCCATCAGG	1320
GAAGAACCCT	ATACTTGGTT	TGCTACCCTT	AGTATTTATT	ACTAACCTCC	CTTAAGCAGC	1380
AACAGCCTAC	AAAGAGATGC	TTGGAGCAAT	CAGAACTTCA	GGTGTGACTC	TAGCAAAGCT	1440
CATCTTTCTG	CCCGGCTACA	TCAGCCTTCA	AGAATCAGAA	GAAAGCCAAG	GTGCTGGACT	1500
GTTACTGACT	TGGATCCCAA	AGCAAGGAGA	TCATTTGGAG	CTCTTGGGTC	AGAGAAAATG	1560
AGAAAGGACA	GAGCCAGCGG	CTCCAACTCC	TTTCAGCCAC	ATGCCCCAGG	CTCTCGCTGC	1620
CCTGTGGACA	GGATGAGGAC	AGAGGGCACA	TGAACAGCTT	GCCAGGGATG	GGCAGCCCAA	1680
CAGCACTTTT	CCTCTTCTAG	ATGGACCCCA	GCATTTAAGT	GACCTTCTGA	TCTTGGGAAA	1740
ACAGCGTCTT	CCTTCTTTAT	CTATAGCAAC	TCATTGGTGG	TAGCCATCAA	GCACTTCCCA	1800
GGATCTGCTC	CAACAGAATA	TTGCTAGGTT	TTGCTACATG	ACGGGTTGTG	AGACTTCTGT	1860
TTGATCACTG	TGAACCAACC	CCCATCTCCC	TAGCCCACCC	$\tt CCCTCCCCAA$	CTCCCTCTCT	1920
GTGCATTTTC	TAAGTGGGAC	ATTCAAAAAA	CTCTCTCCCA	GGACCTCGGA	TGACCATACT	1980
CAGACGTGTG	ACCTCCATAC	TGGGTTAAGG	AAGTATCAGC	ACTAGAAATT	GGGCAGTCTT	2040
AATGTTGAAT	GCTGCTTTCT	GCTTAGTATT	TTTTTGATTC	AAGGCTCAGA	AGGAATGGTG	2100
CGTGGCTTCC	CTGTCCCAGT	TGTGGCAACT	AAACCAATCG	${\tt GTGTGTTCTT}$	GATGCGGGTC	2160
AACATTTCCA	AAAGTGGCTA	GTCCTCACTT	CTAGATCTCA	GCCATTCTAA	CTCATATGTT	2220
CCCAATTACC	AAGGGGTGGC	CGGGCACAGT	GGCTCACGCC	TGTAATCCCA	GCACTTTGAG	2280
AGGCTGAGGT	GGTAGGATCA	CCTGAGGTCA	GGAGTTCAAG	ACCAGCCTGT	CCAACATGGT	2340
GAAACCCCCA	TCTCTACTAA	AAATACCAAA	AATTAGCCGA	GCGTAGTGAC	GGGTGCCCGT	2400
AATCCCAGCT	ACTCAGGAGG	CTGAGACAGG	AGAATCACCT	GAACCCCAGA	GGCAGAGGTT	2460
			CAGCCTGGGC			2520
			AAACAGTCTA			2580
			TGCCTGTAAT			2640
			TGAGACCAGG			2700
			AGCCGGGCAT			2760
			TTGCTTGAGC			2820
TCAGCCCTGA	TTGTGCCAGT	GCACTCCGGC	CTGGGTGACA	GAGTGAGACC	CGTGCTCAAA	2880
			CAGGAGTTTG			2940
			GCGAGACTCC			3000
AAAATTATCT	GAATGATCCT	GTCTCTAAAA	AGAAGCCACA	GAAATGTTTA	AAAACTTCAT	3060
			AGCACTTAAA			3120
			CACATAATTA			3180
AGAATGGCTA	CCCCATTCTC	TAGACAAAAT	CAAATTGTCC	TATTGTGACT	CTTCTAAAAA	3240
TGAAGATGAA	GAGCTATTTA	ATGACACACC	TTGGATTAAA	ACGGGAATCA	CATCTTAAAG	3300
					GACAAGTCTC	3360
					CTTAGGATCG	3420
					GCCCTGAGAC	3480
					TGTCTCACAC	3540
					AGGAGTTCAA	3600
					ATTAGCCGGG	3660
CATGGTGGCA	. GGCGCCTATA	ATCCCAGCTA	CTGGGGAGGC	TGAGGCAGGA	GAATC	3715



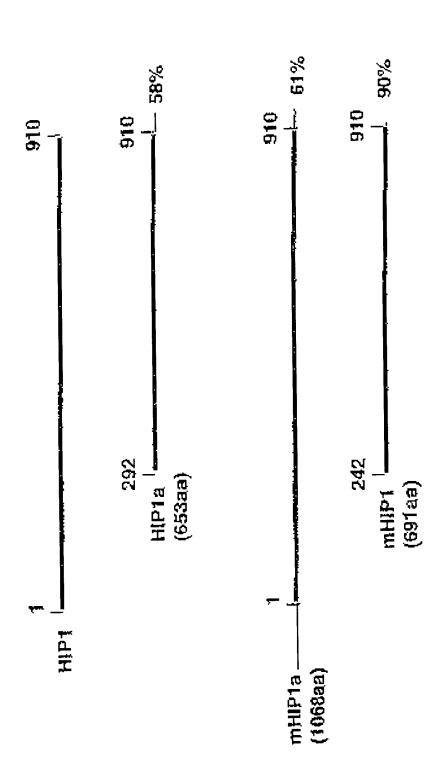


[7] -ep -/

Fy 2. HIP1 Clones: Nucleotide Alignment



جُمْع HIP1 Clones: Protein Alignment



>Usumpin A

SAEVTHOVERALDTDEXKMLLFLCRDVALDVVPPRVNDLIDLLRERGKLSVCDLRELLYRVSRFDIJKRULK

>Usurpin, B

YRVIMAHIGEDLDKSDVSSLIF:MXDYMGRGKISKBKSFLOLVVELSKLVLVAPDQLDLLEKCLKNIHRIDLKYKIQK

PSRNLYDIGH QD5EDLASLKELSI,DYIPQRKOUPIKRALMIYQRIADEKRALEESALSFLKELLFRINRLOLLITYLN

>Casp~8 B

YRVIGYOISEEVEREEREENTLORSISKCKLODOMNLLDIFTEMEKKVILGECKLOILKKVCAQINKSLAXIND

PREXILTIDEN LGVQDVENDKFLOIGLVFRAKKDERSSSASDVFENLLANDLLEESDPFFLAELDYTTRQXXLDQHLNO

>Casp-10: B

FRALLYBLANGI DSBACKDAIFLL KODDIKTEMISLSKY AKUBYGGXI DEDALI CLEDICKTVVPKLLRNIEK

fly(1,85v\$55l\$35fl7flkflcl6ry6rrxjfry05gldlf3h5lfQndlefghtelirbllastlasirbrdllrrydd

>MC159 ሕ

SLPFLRHILDELDSWEDSLUKSKKHORRPGCTTVTQRLCSLSQQRXDTLARLVEMLYVLORMDLLKSRYC

>MC159 B YMKLHYCYGEENDSSELKALRI.FACNLNPSLSTALSESSRPVELYLALENYGLYSSSSYSYLÄDMLRTLSSLDI.CQDLVL

CRCLMALVNDF130X6VEERYF1CAPRIESH12PGSKKSF1R1AS11ED1E11ZGDKIFF1RH11TT1GRADIAVN160V

>K3 orfk13A TYEVICEVARKLG/DDREVVLFLLBVFLFQFFLAQS/GALRALKEEGRS/FFFLLAECLPRAGBRULLRULDS

>KS orfki39

YOLTVLHYDGELCARDIRSLIFLSKDTIGSRSTPOTFLEDVYCHENIDIJCPTDVDALMSDLRSLSRVDIQROVQT

SELERDIABÇONIROQAADDÇEYIRRELDE),ARQREDTEKAQRSISSTERKAQANEQRYSKI,KEKYSELVONBADIJ,RKN Aŝ.

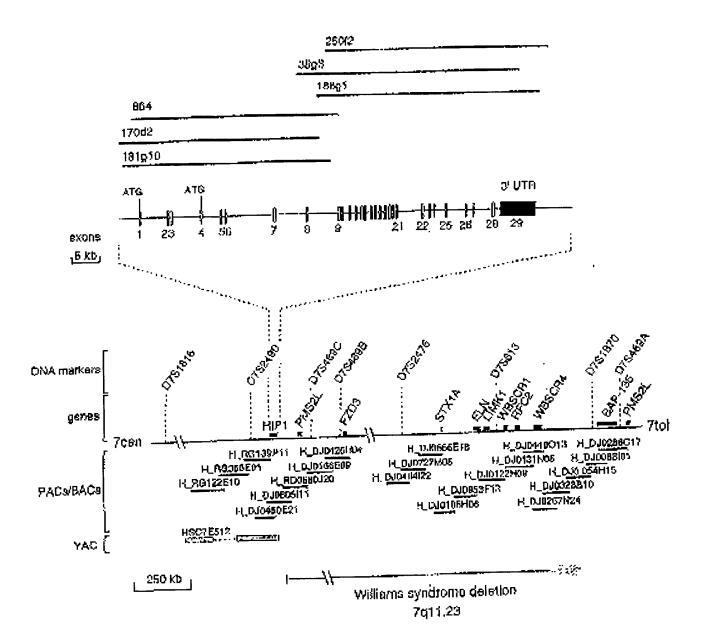
>UIPla

GEDBEORKOKOKALVONEOLRHELAQLUAAQLERERSOGLREEAERKASATEARYNKIJKEKESELVIIVHAELLRIMAD

>mHIPla. nglearleeonnokokalvoneolrhelaqinaldlerarhogireererkasatearyskikekkelijitkabilirka ኢው

14IHm<

SELEAELNEODILGROAMDLCEYLRTFIJDRIKBOREDTEKAGRELTETERKAGRERORYBRIKERYSELVONRADLERM ÆΕ



#19 B

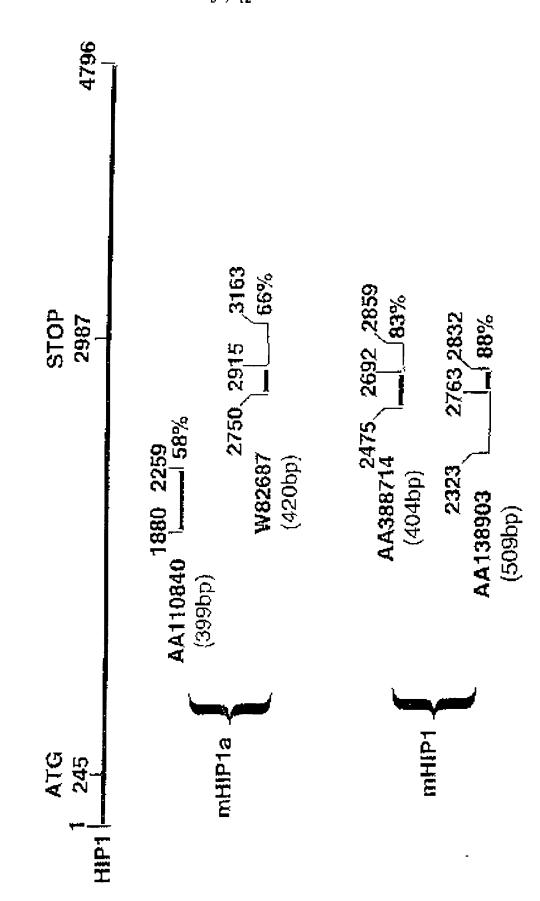
8.6 8.2	160	2 sign 2 5 1	ID 15- 10-15 10-15	4 th 4 th 4 th	0 K P	n -a 	ര <b>ഗ</b> ന ശ ന ൾ	r- 9 1- 9	63Z 751
SORLTHUILISEEL.	RY W S Delected by Structure by Caraman Structure by Stru	W L V. X S. EMPOYLECEUNINGAT VENSIONSREVENSION ENGLEGHENSION VYCARELERHISILE	il to exel a Ilpomendentals Ilpomentals	2 I I SOUTH THE STATE OF THE ST	A . B SREA A B 3 K P Y Achelebraragan eg-atsmirestre- ar	. ರಿಸ್ತುರಿಕ್ಕಾರ್ಪಕ್ಕನ್ನು . ಶಿರಾಸುಗಾರಿಗಳಿಗೆ	ar c c c c c control of c c control of c control of con	מבתאבושה - בסבר ו המבתאבושה - בסבר ו מבמיר - במאפונית (	LEVRE YE WALFAREALECTES LELEVAREALECTES
AGGSERGGUFEDISH GOFLIFELIGISLED.	· Hatledorp v Verktledorproes Lybatledorgenes		. Seklety e. Fyresklotykalig Yerskagirkylvs		r s skra k kanseserakana arverenasian	2 3.68 ONQREATSORES SACRALTHOS	A Q P LACETYCHASTISCENDT A Q P	r Ceraterades Antachesterad	ta ra en electro 1 Alegania (semento) 2 alegania espe
ORCHDACHGISGAN ORCH	L ve hance areasonasca activatersative	ed. n. n. 7 Problem molity Tu-eddeden mente	ER RF 8 E ERCROMBOPTELATEL EGGSTELLFERTERS	o n d medassessessessessessessessessessessessesse	A B B T SHLAQQAADDCERTA ABLDCERRQRADISK KASHDENESRGREK MELAUACNSKICKDO	OR ISDGGGRATJEGLEY KVEKADA	ar Esmoqlakoqrighli Leklevales-Gyli	er s Labelister Actilotterska	di en
444,690,846,690,849,	E EE PA V Bebroagieks vaa Bebrosgefahtvur	e po l loga Glomomomosens Glomosens	28 . E G ER My & K. ESCL?ADYLOG ERCHOMBQFTFLACE 88Q/P	O N O O O O O O O O O O O O O O O O O O		O C CTROYSHARDAQV ÖLSREHKESGDESSER ROZGLERAGUSKE RENESITAKR	ನಿ ನಿಸರಾಭ್ರವಗಳುವಾಣ -೧೮೪೪೦ನಿನಿಸಲಾಗಿನಗಳ	I B G A SECIECASINE INBELSON SECTE CONTROL ACTIONS A TELEVISION OF A CONTROL ACTIONS A TELEVISION A CONTROL ACTIONS AS A CONTROL ACTION	
	r ename e er Kerenarotlot Korenatieves	V ELS BE VERELENS VERTERS VECTOR PROPERTY NAMED IN THE PROPERTY NA	GQC FEET TED S YOUT OF SECTION OF	L F F E KASHIDEVIVIENE ESPOSERVIMENCI IINIDASQUALFONITF DOIPGESENSOPPHF DAGSYRITZHAYLUSE GA	B PHYTES QAYVIQ DYGHVEBREBERGE PLICED RARIEG YZYRILDAMOKETER	E L RQ. LVGNEADJURM ABVIRGUSHARDAQV PRØBSVLANTEL EDROZGLEASJUSKR	-MANAGEREACENT THE BEST RESERVED B H		EP Volskirags <u>olv</u> e Rodigiesskerses
MALOQUSEHRREQQL GTANKGMUFUFQDA	Q KAI B . A KHART I.OT 9202081NEAINTHE ANYESERABUGILOT PAGUSANGRAITHES VENEDREARTILINGS	n el geg mons.eskolulsi namintsolopojes					REARPARICERCE ARLITY-SELECTION		<i>t. v</i> betaslækadstak <u>k</u> fo sodsalskækdigatl. ko
1 hips 2 zk370.3	1 h122 2 ex370.3	1 bipi 2 28370.3	1 hipt 2 skate.3	2 88375.3	1 5ip1 2 2K2JC.5	1 hip: 7 xk370.3	ι μέρι ε 28370.1	1 bigt 2 88570.3	1 bişt 2 2k379.3

# 7 (CONT.)

60 LV 60 - 1 18 10	8) 67 66 63 63 16 61
AD UV CREK 7 % N. V. BIABSTA.	DES L R EM SYN I BLE L ER EL
ADLUTYGEREKF BELLMVCHEIRAGER.	DESSMIL HÖLERGENSSYLVIL KLENNLYGERGE
ADGUTGEGKF BELLTANGELLABTA	LISTEL HÄRKKEDORSONINL BLEZPINGERURLAA
28 EST i hisols a kand a v v on chek.	d vor d . Dea t.
Sek kerandhasis abkandhorta abbuyosak.	Takwasilsexeti Betok-kurskmil
1824 kekronghisols akkanduanturs abbuniosok.	Takwaskaceti Betok-kurskmilt
MANDALINASKOV PRETVERSERES SLASVENGUNISSKU PRETVESKRES SLASVENGUNISSKU GIRUMAGRACI	gi. e vildess 1, i, a. ung in val d
1 tigh	3 hipl
2 = 13 - 1	2 zk370.3

LOK ON A

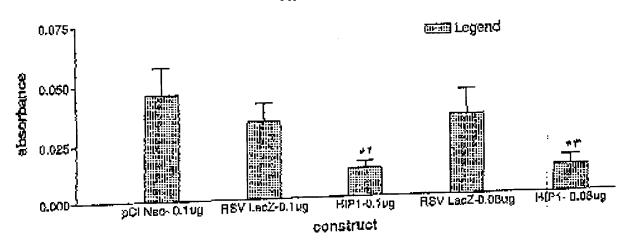
# Fig 7 Mouse ESTs



milldy1.pairs@reph-2 - Tuo Ayr 28 11:30:41 1658

Stress .

HIP1 TOXICITY



Fy 8

HIP1 transfected into HD1955-15 stable cell line 36 hr post-tansfection

Hig-1 in taxic in the presence of hunder

Absorbance
0.15
Absorbance
0.00
HIP1 Lec2 Vestor Mock

Transfected constructs

Fug 94

HIP1 transfected into HD1955-128 stable cell line 36 hr post-tansfection

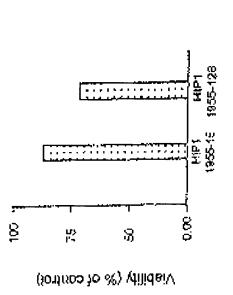
Hip-1 is toxic in the presence of hushing

0.15 A C.05 HIP1 Lac2 Vector Mock

Transfected constructs

F. 298

Polyglutamine-dependence of HIP-1 taxicity



Transfected constructs/cell lines

798

# SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: Kalchman, Michael

Hayden, Michael R.

Hackam, Abigail

Chopra, Vikramjit Singh

Nicholson, Donald W.

Vallaincourt, John P.

Rasper, Dita M.

(ii) TITLE OF INVENTION: Apoptosis Modulators That Interact with the

Huntington's Disease Gene

- (iii) NUMBER OF SEQUENCES: 44
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Oppedahl & Larson
- (B) STREET: PO Box 5270
- (C) CITY: Frisco
- (D) STATE: CO
- (E) COUNTRY: USA
- (F) ZIP: 80443-5270
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: MS DOS 5.0
- (D) SOFTWARE: WordPerfect
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Larson, Marina T.
- (B) REGISTRATION NUMBER: 32038
- (C) REFERENCE/DOCKET NUMBER; UBC.P-013US2
- (jx) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (970) 668-2050
- (B) TELEFAX: (970) 668-2052
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) \$BQUENCE CHARACTERISTICS:
- (A) LENGTH: 1164
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SBNSE: no



- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: cDNA for Huntingtin-interacting protein
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACAGCTGACA	$\tt CCC'IGCAAGG$	CCACCGGGAC	CGCTTCATGG	AGCAGTTTAC	50
aaagttgaaa	CATCTGTTCT	ACCCCTCCAC	CAACCTCCAG	TACTICAAGC	100
GGGTCATTCA	GATCCCCCAC	CTCCCTGAGA	ACCCACCCAA	CTTCCTGCGA	150
GCCTCAGCCC	TGTCAGAACA	TATCAGCCCT	${\tt GTGGTGGTGA}$	TCCCTGCAGA	200
GGCCTCATCC	CCCGACAGCG	AGCCAGTCCT	AGAGAAGGAT	${\tt GACCTCATGG}$	250
ACATGGATCC	CTCTCAGCAG	$\mathbf{NMM}_{\mathbf{M}}\mathbf{MM}\mathbf{M}\mathbf{M}\mathbf{M}$	ACAACAAGTT	TGATGACNTC	300
TTTGGCAGTT	CATCCAGCAG	TGATCCCTTC	AATTTTCAACA	GTCAAAATGG	350
TGTGAACAAG	GATGAGAAGC	ACCAUTTAAT	TUAUCGAUTA	TACAGAGAGA	400
TCAGTGGATT	GAAGGCACAC	CTAGAAAACA	TGAAGACTCA	GAGCCAGCGG	450
GTTGTGCTGC	AGCTGAAGGG	CCACGTCAGC	GAGCTGGAAG	CAGATCTGGC	500
CGAGCAGCAG	CACCTGCGGC	AGCAGGGGGC	CGACGACTGT	GAATTCCTGC	550
GGGCAGAACT	GGACGAGCTC	AGGNGGCAGC	GGGAGGACAC	CGAGAAGGCT	600
CAGCGGAGCC	TGTCTGAGAT	AGAAAGGAAA	GCTCAAGCCA	ATGAACAGCG	650
ATATAGCAAG	CTAAAGGAGA	ACTACAGCGA	CCTCCTTCAG	AACCACCCTG	7D0
ACCTGCTGCG	GAAGAATGCA	CAGGTGACCA	AACAGGTGTC	CATGGCCAGA	750
CAAGCCCAGG	TAGATTTGGA	ACGAGAGAAA	AAAGAGCTGG	AGGATTCGTT	800
GGAGCGCATC	AGTGACCAGG	GCCAGCGGAA	GACTCAAGAA	CAGCTGGAAG	850
TTCTAGAGAG	CTTGAAGCAG	GAACTTGGCA	CAAGCCAACG	GGAGCTTCAG	900
CTTCTGCAAG	GCAGCCTGGA	AACTTCTGCC	CAGTCAGAAG	CAAACTGGGC	950
AGCCGAGTTC	GCCGAGCTAG	AGAAGGAGCG	GGNCNGCCTG	GTGAGTGGCG	1000
CAGCTCATAG	GGYGGYAY	TTATCTGCTC	TTCGGAAAGA	ACTGCAGGAC	1.050
ACTCAGCTCA	AACTGCCCAG	CACAGAGGAA	TCTATGTGCC	AGCTTGCCAA	1100
AGACCAACGA	AAAATGCTTC	TGGTGGGGTC	CAGGAAGGCT	GCGCACCAGG	1150
TGATACAAGA	CGCG				1164

- (2) INFORMATION FOR SEQ ID NO:2:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 386 (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL, no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: Huntingtin-interacting protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Phe Thr Lys Leu Lys Asp Leu Pho Tyr Arg Ser Ser Asn Leu Gln
20 35

Tyr Phe Lys Arg Val Ile Gln 1le Pro Gln Leu Pro Glu Asn Pro

,	WO 99	<b>6098</b> 6			)								РСТЛ	)S99/117
	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	<b>V</b> V/=		35					40					45
Pro	Asn	Phe	Leu	Arg 50	Alu	Ser	Ala	Leu	5er 55	Ģlu	His	Ile	Ser	Pro 60
νal	Va1	Va?	Ile	Pro 65	Ala	Clu	Ala	šer	Ser 70	Pro	Asp	Ser	Glu	Pro 75
Val	ьеu	Glu	IJя	Asp 80	Asp	Leu	Met	Asp	<b>Met</b> 85	Asp	Ala	Ser	Glo	Gln 90
Asn	Leu	Phe	Asp	<b>Л</b> вп 95	Lys	Phe	Asp	Ąsp	Phe 100	Glγ	Ser	Ser	Ser	Ser 105
Ser	Asp	Pro	Брь	Asn 110	Phe	Asn	ser	alD	Asn 115	Cly	Val	Asn	Lys	Asp 120
Glu	Lys	Asp	His	Leu 125	Ile	Glu	Arg	Leu	Туг 130	Arg	Glu	Ile	Ser	Gly 135
Leu	Lув	Ala	Gln	Leu 140	Glu	Asn	met	Ъуs	Thr 145	Glu	Ser	Gln	yrg	Vall. 150
yal	Leu	Cln	Leu	Lys 155	ĜĴУ	Bis	Val	Ser	Glu 160	Leu	Glu	Ala	Asp	Leo 165
λla	GJ.11	Gln	Gln	His 170	Leu	Arg	Gln	Gln	λla 175	λla	Asp	Asp	Cys	G)u 180
Phe	Ъeu	Arg	Ala	GJ 12 1.85	L€u	Asp	Glu	LCu	Arg 190	Cln	Arg	Glu	Asp	Thr 195
G) 11	Lys	Ala	Gln	Arg 200		Leu	Ser	Glu	Ile 205	Glv	Arg	Турв	Ala	Gln 210
Ala	. Asn	Glu	Gln	Arg 215		Ser	Lys	Leu	Lys 220	Glu	Lys	Туг	Ser	61u 225
Leu	. Val	Gli	Asn	His 230		yst	Lev	. Leu	Arg 235	ĮŅĖ	Аал	Ala	Glu	Val 240
Thr	Lys	Glr.	Val	Ser. 245		. Als	Arg	3 Gln	Ala 250	Gln	val	Asp	) beu	Glu 255
Arg		ı Lyr	ı Lye	Gln: 260		ı Glı	ı Asp	Ser	ъец 265		λrg	I1∈	: Ser	Asp 270
		/ Clr	Arg	руя 275		с G:l.т	1 G).1	ı Glin	280	Glu	.Val	Leu	: Clu	ser 2B5
Lev	a Liyi	g Glı	n Glu	LCL	1 G37	y 'lh:	r Sei	r Glx	) Arg	Glu	Len	Gln	ı Val	. Len

WO 99/60986	PCT/US99/11743

290   295   300   301   301   305   307   307   310   310   310   310   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315   315		4441.22	OURANG	_							_			100	COLMITIONS
Ala Glu Phe Ala Glu Lou Glu Lys Glu Arg Asp Ser Leu Val Ser 320  Gly Ala Ala His Arg Glu Glu Glu Leu Ser Ala Lou Arg Lys Glu 345  Leu Gln Asp Thr Cln Lou Lys Leu Ala Ser Thr Glu Glu Ser Met 350  Cys Gln Leu Ala Lys Asp Gln Arg Lys Met Leu Leu Val Cly Scr 375  Arg Lys Ala Ala Glu Gln Val Ile Gln Asp Ala					290					295					300
Gly Ala Ala His Arg Glu Glu Glu Leu Ser Ala Leu Arg Lys Glu 345  Leu Gln Asp Thr Gln Leu Lys Leu Ala Ser Thr Glu Glu Ser Met 350  Cys Gln Leu Ala Lys Asp Gln Arg Lys Met Leu Leu Val Gly Ser 375  Arg Lys Ala Ala Glu Gln Val Ile Gln Asp Ala	Gln	Gly	Ser	Leu		Thr	Ser	sfA	Cln		Glu	Ala	Asn	Trp	
Leu Gln Asp Thr Gln Leu Lys Leu Ala Ser Thr Gln Gln Ser Met 350  Cys Gln Leu Ala Lys Asp Gln Arg Tys Met Leu Leu Val Cly Ser 365  Arg Lys Ala Ala Glu Gln Val Ile Gln Asp Ala	Ala	<b>Gl</b> u	Phe	Ala		ьси	Glu	Lys	Glu		qaA	Ser	Leυ	Val	
Cys Gln Leu Ala Lys Asp Gln Arg Lys Met Leu Leu val Cly Scr 365 370 375  Arg Lys Ala Ala Glu Gln Val Ile Gln Asp Ala	G.l Y	Ala	Ala	Ні. в	Arg 335	Glu	G) 13	Glu	Leu		Ala	Leu	Arg	Lys	
365 370 375  Arg Lys Ala Ala Glu Gln Val Ile Gln Asp Ala	Leu	Gln	Asp	Thr		Leu	Lys	Leu	Ala	_	Thr	Glæ	Glu	Ser	
	Сув	Gln	Leu	Ala		Asp	Gln	Arg	Гув		Leu	Leu	Val	Cly	
	Arg	Lys	Ala	Ala		Gln	Val	Ile	Gln						

# (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4796
- (B) TYPB: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: cDNA for Huntingtin-interacting protein
- (xi)SEQUENCE DESCRIPTION; SEQ ID NO; 3;

CAGTGTACGG	TTGATCATAT	AACGCCCCCC	GCCGCCATTC	GTTTATATAT	50
CGCAAATTGA	${\bf TNTAGGGGGG}$	GGGGGATGGN	CAGAGATTTC	GCTTCATTAG	100
GCCATTATAA	${\tt GCAGGAAGGG}$	TTTCAAGGAA	AAAAACCCAG	AAAGTGCATA	150
TTGCACCCAC	CATGAGAAAG	GGGCAACAGA	CCTTNTGTTN	TGTTNTCAAC	200
CGCCTGCTTC	TGTTTTAGCA	ACGCAGTGTT	TTGGTCCAAC	TTGTCCCATG	250
			CGAACGTCCT		300
GTGAGATACA	GAAATGAATT	GAGTGACATG	AGCAGGATGT	GGGGCCACCT	350
GAGCGAGGGG	TATGGCCAGC	TGTGCAGCAT	CTACCTGAAA	CTGCTAAGAA	400
CCAAGATGGA	GTACCACACC	AAAAATCCCA	GGTTCCCAGG	CAACCTGCAG	450
			GAMAGTGACG		500
TTTCCAGTTA	ACAGTGGAGA	TGTTTGACTA	CCTGGACTGT	GAACTCAACC	550
TCTTCCAAAC	AGTATTCAAC	${\tt TCCCTGGACA}$	TGTCCCGCTC	TGTGTCCGTG	600
ACGGCAGCAG	GGCAGTGCCG	$\tt CCTCGCCCCG$	CTGATCCAGG	TCATCTTGGA	650
CTGCAGCCAC	CTTTATGACT	ACACTGTCAA	GCTTCTCTTC	AAACTCCACT	700
CCTGCCTCCC	AGCTGACACC	CTGCAAGGCC	ACCECCACCE	CTTCATGGAG	750

WO 99/60986 PCT/DS99/11743

7,42					000
			ADDADOTODO		800
			CCCTGAGAAC		850 040
102100000	CTCAGCCCTG			GGTGGTGATC	900
• • • • • • • • • • • • • • • • • • •	CCTCATCCC		CCAGTCCTAG	AGAAGGATGA	950
CCTCATGGAC	ATGGATGCCT		TATTYTTATT	AACAAGTTTG	1000
ATGACATCTT	TGGCAGTTCA	TTCAGCAGTG		TTTCAACAGT	1050
CAAAATGGTG	TGAACAAGGA	TGAGAAGGAC		ACCGACTATA	3100
CAGACAGATO	AGTGGATTGA	AGGCACAGCT	=	AAGACTGAGA	1150
GCCACCGGGT	'I'GTGCTGCAG	CTGAAGGGCC	ACCTCAGCGA		1200
GATCTCCCCG	AGCAGCAGCA	CCTGCGCCAG	CAGGCGGCCG	ACGACTGTGA	1250
ATTUCTGCGG	GCAGAACTGG	ACGAGCTCAG		GAGGACACCG	1300
AGAAGGCTCA	GCGGAGCCTC	TUTGAGATAG	AAAGGAAAGC	TCAAGCCAAT	1350
OTT 10KING A		AAAGGAGAAG		TGGTTCAGAA	1400
			GGTGACCAAA		1450
			GAGAGAAAA		1500
GATTCGTTCG	ACCCCATCAG	TGACCAGGCC	CAGCGGAAGA	CTCAAGAACA	1550
GCTGGAACTT	CTAGAGAGCT	TGAACCAGGA	${\tt ACTTGGCACA}$	AGCCAACGGG	<b>3600</b>
AGCTTCAGGT	TUTGCAAGGC	AGCCTGGAAA	CTTCTGCCCA	GTCAGAAGCA	1650
AACTGGGCAG	CCGAGTTCGC	CGAGCTAGAG	AAGGAGCGGG	ACAGCCTGGT'	1700
GAGTGGCGCA	GCTCATAGGG	AGGAGGAATT	ATCTGCTCTT	CGGNAAGAAC	1750
TGCAGGACAC		CTGGCCAGCA	CAGAGGAATC	TATGTGCCAC	1800
CTTGCCAAAG	ACCAACGAAA	AATGCTTCTG	GTGGGGTCCA	CCAACGCTGC	1850
GGAGCNGGTG	ΛΊΆCλλGΆCG		<b>GCTTCAACAA</b>	CCTCCTCTCA	1900
TCAGCTGCGC	TGGGTCTGCA	GATCACCTCC	TCTCCACGGT	CACATCCATT	1950
TCCAGCTGCA		GGAGAAAAGC	TGGAGCCAGT	ATCTGGCCTG	3000
CCCAGAAGAC	ATCAGTGGAC	TTCTCCATTC	CATANCECTG	CTGGCCCACT	2050
IGACCAGCGA			CCACCTGCCT	CAGAGCCCCA	2100
CCTGAGCCTG	CCGACTCACT		TGTAACCAGT	ATGCCAGGGA	2150
AACCCTCGCC			AGAGGGAAGC	CTTGAGAATG	2200
CCGACAGCAC	AGCCATGAGG		GCAACATCAA	CCCCATCCCC	2250
GAGGACCTCC	TGCCCAGGGG		AACCAGGAGG	ACCTGGGGGA	2300
CCTGGTGGAC	AAGGAGATGG		AGCTGCTATT	GAAACTTGCA	2350
CGGCCAGAAT		CTCAGCAAAT		AGACACAGGA	2400
	_		CGTTGCTGTA		2450
CACAGAMATAGG	- VAGATOWNIAW	######################################	TAAGGACCTC	CAGAGAGAGA	2500
GUARGUTATT	- C100G1GC1C21	ACAGCA9CCC	CTAAAGAGTT	ביוטויטוים אויים	2550
				CTGTGGGCTG	2600
				CAAGGCAGAG	2650
				TGCTAGCACA	2700
GGAAATTTGA				VCAGCCCCVV	2750
GCCCAGCTTG				GCCACTGCCG	2800
					2850
				AGAGAGAGAG	2900
				GCCAAGAGAT	2950
				CAGAAGGAGC	2950 3000
	GGGAGAGCTT	=		TGCTGGTGTT	
	CCGAAGAACG			CACTGCAAGA	3050
				TATGTCAGTG	3100
TAAATCCTTG	TTACCTATCT	CCTCTCTCTT	ATTTCCCCAG	CCACAGGCCA	3150
				CCCAGTGCCG	3200
AGGACATGCA	TGACACTTCC	CONTAGARAN	TCCATAGCGA	CACCCTTTCT	3250
GTTTGGACCC	: ATGGTCATCT	, Сабатстта	, CCCCCCLLCCC	TAGTTAGCAT	3300

WO 99/6098	96			PCT	70899/11743
CCAGGCTGGC	CAGTGCTGCC	CATGACCAAG	CCTAGGTACG	AAGAGGGGTG	3350
GTGGGGGCA	GGGCCACTCA	ACAGAGAGGA	CCAACATCCA	GTCCTGCTGA	3400
CTATTIGACC	CCCACAACAA	TEGGTATOOT	TAATAGAGGA	GCTGCTTGTT	3450
GTTTGTTGAC	AGCTTGGAAA	GCGAAGATCT	TATGCCTTTT	CTTTTCTGTT	3500
TTCTTCTCAG	TCTTTTCAGT	TTCATCATTT	$\tt GCACAAACTT$	GTGAGCATCA	3550
GAGGGCTCAT	CGATTCCAAA	CCAGGACACT	${\tt ACCCTGAGAT}$	CTGCACAGTC	3600
AGAAGGACGG	CAGGACTOTO	CTGGCTGTGA	${\tt ATGCCAAAGC}$	CATTCTCCCC	365 <b>0</b>
CTCTTTGGGC	AGTGCCATGG	ATTTCCACTG	${\tt CTTCTTATGG}$	TGGTTGGTTG	3700
GCTTTTTTGG	<del>ተ</del> ስተተርታዊ ተ	TTTTTTTAAG	TTTCACTCAC	ATAGCCAACT	3750
CTCCCAAAGG	GCACACCCCT	GGGGCTGACT	CTCCAGGGCC	CCCCAACTGT	3800
<b>GGTAGCTCCA</b>	GCGATGGTGC	TGCCCAGGCC	${\tt TCTCGGTGCT}$	CCATCTCCGC	3850
CTCCACACTG	ACCAAGTGCT	GGCCCACCCA	${\tt GTCCATGCTC}$	CAGGGTCAGG	3900
CGGAGCTGCT	GAGTGACAGC	TTTCCTCAAA	AAGCAGAAGG	AGAGTGAGTG	3950
CCTTTCCCTC	CTAAAGCTGA	ATCCCGGCGG	AAAGCCTCTG	TCCGCCTTTA	4D00
CAAGGGAGAA	GACAAÇAGAA	AGAGGGACAA	GAGGGTTCAC	ACAGCCCAGT	4050
TCCCGTGACG	AGGCTCAAAA	ACTTGATCAC	ATCCTTGAAT	GGACCTCCTC	4100
AGATCAACAA	CACTACTTCC	CTGCCGGAAT	GAACTGTCCC	TCAATGOTCT	4150
CTGTCAAGCG	GCCCGTCTCC	C1"TGGCCCAC	${\tt AGACGGAGTG}$	TGGGAGTGAT	4200
TCCCAACTCC	TTTCTCCACA	CGTCTGCCTT	$\tt GGCATCCTCT$	TGAATAGGAA	4250
GATCGTTCCA	CTTTCTACGC	AATTGACAAA	CCCGGAAGAT	CAGATGCAAT	430C
TGCTCCCATC	AGGGAAGAAC	CCTATACTTG	GTTTGCTACC	CTTAGTATTT	4350
ATTACTAACC	TCCCTTAAGC	AGCAACAGCC	TACAAAGAGA	TCCTTCGACC	4400
AATCAGAACT	TCAGGTGTGA	CTCTAGCAAA	GCTCATCTTT	CTGCCCGCCT	4450
ACATCAGCCT	TCAAGAATCA	GAMGAMAGCC	AAGGTGCTCC	ACTGTTACTG	4500
<b>ACTIGGAT</b> CC	CNANGCANGG	AGATCATTTG	GASCICITCG	CTCAGAGAAA	4550
ATGAGAAAGC	ACAGAGCCAG	CCCCTCCAAC	TCCTTTCAGC	CACATGCCCC	4600
AGGCTCTCGC	TGCCCTGTGG	ACAGGATGAG	${\tt GACAGAGGGC}$	ACATGAAÇAG	4650
CTTGCCAGGG	ATGGGCAGCC	CAACAGCACT	${\tt TTTCCTCTTC}$	TAGATGGACC	4700
CCAGCATTTA	AGTGACCTTC	TGATCTTGGG	AAAACAGCGT	CTTCCTTCTT	4750
TATCTATAGC	AACTCATTGG	TEGTAGCCAT	CAAGCACTTC	GGAA'I'I'	4796

### (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 924
- (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: no.
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FBATURE: Huntinglin-interacting protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Ser Arg Met Trp Gly His Lew Ser Glu Gly Tyr Gly Gln Lew 1 5 10 15

Cys Ser Ile Tyr Leu Lys Leu Leu Arg Thr Lys Met Glu Tyr His 20 25 30

												1		
•	WO 99	<b>1609</b> 86			•								PCTA	J <b>8</b> 99/11743
Thr	Lys	λsn	Pro	Arg 35	Phe	Pro	Gly	Авп	Leu 40	Gln	Met	Ser	Asp	Arg 45
Gln	Leu	Asp	Glu	<b>Al</b> a 50	Cly	Glu	Ser	Asp	Val 55	Asn	Agn	Phe	Phe	Gln 60
Leu	Thr	Val	G1.v.	Мет. 65	Phe	Asp	Tyr	Tæp	Glu 70	Сув	Glu	Leu	λsn	Leu 75
Phe	Gln	Thr	Val	Phe 80	Asn	Ser	Len	ysb	Met B5	Ser	Arg	ser	Val	Ser 90
Val	Thir.	Ala	Ala	G3y 95	Gln	Сув	Arg	Leu	Ala 100	Pro	Leu	Il.e	Gln	Val 105
Ile	Leu	Asp	Cys	Ser 110	llis	Leu	Тух	Asp	Tyr 115	Thr	Val	Lys	Leu	Leu 120
Phe	Lув	Leu	His	Ser 125	Cys	Leu	Pro	Ala	Asp 130	Thr	Len	Gln	GJA	Нів 135
Arg	Asp	Arg	Phe	Met 140	Glu	Gln	Phe	Thr	Lys 145	Leu	Lys	Asp	Leu	Phe 150
Tyr	Arg	Ser	Ser	Asn 155	Leu	Gln	Туг	Phe	Lys 160	Arg	Leu	Ile	Glu	11e 165
Pro	Gln	Беи	Pro	GJ.u 170	Agn	Pro	Pro	Asn	Phe 175	Leu	Arg	Ala	Ser	Ala 180
LCΩ	Ser	Glu	His	Ile 185	Ser	Pro	Val	Va1	Va1 190	lle	Pro	λla	Glu	λ1a 195
Ser	Ser	Pro	Asp	Ser 200	Glu	Pro	Val.	Leu	Gl u 205	Lys	Asp	Asp	Leu	Met 210
Asp	Met	Asp	s (A	Ser 215	Cln	Cln	Asn	Leu	Pho 220		Asn	Lys	Phe	Asp 225
Asp	Ile	Ph∺	Gly	5er 230		Phe	: Ser	Ser	Asp 235		Phe	Agn	Phe	Авл 240
Ser	Gln	Asn	Gly	Val 245		ГАЗ	: Asp	Glu	Lys 250		His	Leu	Ile	Glu 255
Arg	ьеи	Тук	• Arg	Glu 260		: Ser	Gly	Leu	. <b>Гу</b> я 265		Gln	Leu	Glu	Asn 270
Met	Lγs	Thi	: Glu	Ser 275		Arg	y Val	. Val	. Leu 280	Gln	Leu	Lys	cly	Ris 285

								(					
	99/6 <b>098</b>												US99/11743
Val Se	r Glu	Ъeu	G1u 290	Ala	ysb	ь <del>е</del> и	λla	Glu 295	Gln	Cln	His	Leu	Arg 300
Glzı Gl	afA n	Ala	Asp 305	Asp	Сув	Clu	Phe	Leu 310	Arg	Ala	Glu	Leu	Asp 315
Glu Le	u Arg	Arg	Gln 320	Arg	Glu	Asp	Thr	Glu 325	Lys	Ala	Gln	Arg	Ser 330
Leu Sc	r Glu	Ile	G1u 335	Arg	Lys	λla	Gln	λ1a 340	Asn	Glu	Gln	Arg	' <b>Tyr</b> 3 <b>4</b> 5
Ser Ly	a Leu	Lys	Glu 350	Lys	Тут	Ser	Glu	Leu 355	Val	Gln	Asn	His	Ala 360
Asp Le	น Leu	Arg	Lув 365	Aan	Ala	ĠĴij	Val	ТЪт 370	Iva	GЛп	va I	Ser	Met 375
Ala Ar	g Gln	Ala	Gln 380	Val	Asp	Leu	Glu	Arg 385	Glu	Lys	Lys	Glu	Leu 390
Glu As	p Ser	Leu	Glu 395	Arg	T.l.e	Ser	qaA	Gln 400	Gly	Gln	Arg	Lys	Thr 405
Gln Cl	u Gln	Leu	Glu 410	Val	Lew	Glu	Ser	Leu 415	Lys	Gl'n	Glu	Leu	Gly 420
'I'hr Se	r Glu	Arg	Glu 425	Leu	Gln	lsV	Leu	Gln 430	Gly	Ser	Leu	Glu	Thr 435
Ser Al	a Cln	Ser	Glu 44D	λla	Asn	Trp	λla	λla 445	Glυ	Phe	λla	Glu	Leu 450
Glo Jay	ន Glu	Arg	<b>Asp</b> <b>45</b> 5	Ser	Leu	Val	Ser	Gly 460	Ala	Ala	His	Arg	Glu 465
Glu Gl	u Leu	Ser	A1a 470	Leu	Ang	Lys	Glu	Leu 475	αLĐ	Asp	ጥኪጕ	GJ.m	ьер 480
Lys Le	u Ala	Ser	Thir 485	Glu	Glu	Ser	Met	Сув 490	Gln	Leu	Ala	Lys	Авр 495
Gln Ar	g Lys	мet	Նeu 500	Leu	<b>V</b> al	Gly	Ser	λrg 505	Lys	sΓΛ	Λla	G1u	Gln 510
Val I)	e Gln	Asp)	Ala 515	Leu	asa	Gln	Leu	ցյո 520	Gјп	Pro	ΡĽΟ	Leu	11e 525
Ser Cy	s Ala	Gly	Ser 530	Ala	Asp	His	Leu	Leu 535	Ser	Thr	Val	Jihr	Ser 540

												1	PCTA	(IS99/11743
	WO 99 -			-7.	/17 - a	01.	T	<b>01.</b>	r a cou	r.o.o.	Meres	Car		
lle	Ser	Ser	Cys	545	GIU	Q I D	Leu	GIU	550	261.	Trþ	рет	( <del>3</del> .111	555
Leu	Ala	Сув	Pro	G1u 560	qaA	Tle	Ser	Glγ	Lev 565	l'e/i	Hi.s	Sex	Ile	Thr 570
Leu	Leu	Ala	Hìs	ьеи 5 <b>75</b>	Thr	Ser	Asp	Alæ	11e 580	Ala	His	G3y	Ala	Thr 585
Thr	Cys	Leu	Aıy	Ala 590	Pro	Pro	Glu	Pro	Ala 595	Asp	Ser	Leu	ΤΊλτ΄	Glu 600
Ala	Cys	гав	Gln	<b>Ty</b> r 605	G.l.y	Ary	Glu	Tbr	Leu 610	A).a	<b>ፐ</b> ሃ1 ⁻	Leu	Ala	Ser 615
Leu	Glo	Glu	Glu	Gly 620	Ser	Leu	Glu	<b>A</b> sn	Ala 625	Asp	Ser	Thr	Ala	<b>м</b> еL 630
Arg	Asn	Cys	Lev	Ser 635	Lys	Ile	Ly≊	Alä	11e 640	Gly	Glu	Glu	Leu	Leu 645
Рхю	Arg	Gly	Leu	Asp 650	Ile	Lys	Glin	Glu	Glu 655	Leu	Gly	Asp	Leu	Val 660
Asp	Lys	G1u	Met	Ala 665	Ala	<u> </u>	Ser	Ala	Ala 670	Ile	Glu	Thr	Суѕ	T'inr 675
Alu	Arg	Ile	Glu	Glu 680	Met	Leu	Ser	Iys	Ser 685	Arg	Ala	Gly	Aep	Thr 690
G1y	Val	Lys	. Leu	Glu 695	val	Asn	Glu	hrg	11e 700	Leu	Arg	Cys	Cys	Thr 705
Ser	Leu	Mel	Gln	Ala 710		Gln	. Val	Leu	Ile 715		Ala	Ser	J _i ya	Авр 720
Lev	G.lm	Arg	; Glu	11e 725		Glu	ser	СІУ	730		Thr	Ala	Ser	Pro 735
Lys	: Glv	Ph•	э Туг	740		Ası	Sex	Arg	745	Τίλπ.	· Glu	GJ.y	Len	750
Ser	Ala	s Sei	Lys	Ala 765		Gly	Trp	Gly	770	נו <b>ד</b> ו	· Val	Met	val	. Asp 775
∌J.¥	a Ala	Asj	, Leu	780		. Glr	Gly	, Arc	9 Gly 785	, Lys	: Phe	: Glu	Giv	Leu 790
Met	. Val	l Cy:	s Sei	79:		: Ile	s£A ∈	a Ala	80(	rllT )	. Ale	Clr	ı Ler	Val 805

	<b>ጭ</b> ለ 99	760986											PCT/	US99/11743
λla	Ala	Ser	Lys	val 810	Lys	Ala	Asp	ьуз	Asp 815	Ser	Pro	Asn	Leu	λ1a 820
Gln	Leu	Gln	Gjn	Ala 825	Ser	Arg	Cly	Val	Asn 830	Cln	Ala	Thr	Ala	Gly 835
Val	Val	Ala	Ser	Thr 840	Ile	Ser	Gly	Lys	Ser 845	Glp	ηle	Glu	Glu	Thr 850
Asp	Asn	Net.	qaA	Phe 855	Ser	Ser	Met	Thr	Leu B6D	Thr	Gln	lle	ьуз	Arg <b>8</b> 65
Gln	Glu	Met	Asp	Ser 870	Gln	Val.	Arg	Val	Len 875	Glu	Len	Glu	Asn	Glu B80
Leu	Gln	Lys	Clu	Arg 885	Gln	Lys	Leu	Gly	G1 u B90	Leu	Arg	Lys	Lys	Bi.s B95

Tyr Clu beu Ala Gly Val Ala Glu Gly Trp Glu Glu Gly Thr Glu

Ala Ser Pro Pro Thr Lew Gln Glu Val Val Thr Glu Lya Glu

905

920

910

924

(2) INFORMATION FOR SEQ ID NO: 5

900

915

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1090

(B) TYPE: protein

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: protein

(iii) HYPOTHETICAL: no.

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(ix) FEATURE: Huntinglin-interacting protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met Leu Leu Cys Gln Gly Ser Glu Tmp Arg Arg Asp Gln Gln Leu 5 10 15

Gly Thr Ala Asn Ala Arg Cln Trp Cys Pro Leu Pro Gln Asp Ala 20 25 30

Gln Pro Ala Gly Ser Trp Giu Arg Cys Pro Pro Leu Pro Pro Ala 35 40 45

Gly Arg Leu Gln Gly Thr Asp His Pro Trp Gly Trp Gly Arg Leu 50 55 60

												)		
•	WO 99.	/609R6			•								PC17C	S99/11743
Λla	Gly	G1y	Glу	G).u 65	Arg	Gly	Gly	Leu	Trp 70	Glu	Gly	Leu	Ser	Ні.в 75
Ser	Gln	Arg	Len	lle BO	nis	Leu	Tle	Leu	Leu 85	Ser	Leu	PYO	Len	Leu 90
Val	Phe	Ģ l.π	Thr	val 95	Ser,	Ile	Asn	Lys	Ala 100	lle	Agn	<b>ፒ</b> ኴ፡-	ĞĴπ	Glu 105
Val	Ala	Val	Lув	G)u 110	lys	His	Ala	Arg	Thr 115	Cys	Ile	Leu	Gly	<b>Th</b> r 120
His	His	Glu	Lys	01y 125	Ala	Cln	uhr	Phe	Trp 130	Ser	Val	Val	Asn	Arg 135
Leu	Pro	Leu	Ser	Ser 140	авп	Ala	Va.l	Leu	Cys 145	Trp	Lys	Phe	Cys	мів 150
val	Phe	His	Lys	Leu 155	Leu	yrg	Asp	Gly	Mis 160	Pro	Asn	val	Leu	Lys 165
Vsb	Ser	Tæn	Arg	Туг 3,70	Ar'g	Agn	Glu	Le:12	Ser 175	qań	Met	Ser	Arg	Met 180
Trp	Gly	His	ьеи	Ser 185	Glu	Gly	Тут	Gly	01n 190	ГGл	Сув	Ser	Ile	Tyr 195
Leu	Irys	Leu	Leu	Arg 200	Thir	Lys	Met	Glu	Тух 205		Thx	Гуг	Asn	Pro 210
Arg	Phe	Pro	Gly	Asn 215		Gln	Met.	Ser	Asp 220		Gln	Leu	Asp	<b>Glu</b> 2 <b>2</b> 5
λla	Glγ	Glu	Ser	Asp 230		Asn	Aan	Phe	Phe 235		Leu	Tìnı	Val	Glu 240
Met	. Phe	yeh	Tyr	Նես 245		Cys	Glu	Leu	Asn 250		Phe	Gln	'f'hr	Val 255
₽'n∈	Asn	. Ser	Leu	Asp 260		. Ser	Arg	Ser	val 265		· val	Thr	Ala	Ala 270
Ģly	Gln	Су:	arg	; Leu 275		Pro	Leu	11e	: Cln 288		. lle	Leu	ı Asp	<b>Cys</b> 285
Ser	His	: Le	э Тух	Asp 290		- ጥኬי	r Val	l Lys	Leu 295		. Phe	: Гув	: Leu	His 30D
5è≥	с Суя	: Lei	) Pro	305		כמנוי <	c Lei	2 Clr	31( 31)		AY9	Asp	Arg	Phe 31.5

WO 99/60986		<u> </u>	PCT/US99/11743
Met Gla Gln Pho	e Thr Lys Le	u Lys Asp Leu Phe Tyr	Arg Ser Ser
	320	325	330
Asn Leu Gln Ty	Phe Lys Arg	g Leu Ile Glm Ile Pro 340	Gln Leu Pro 345
Glu Asn Pro Pri	350 Asn Phe Le	u Arg Ala Ser Ala Leu 355	Ser Glu His 360
Ile Ser Pro Val	. Val Val Ile	e Pro Ala Glu Ala Ser	Ser Pro Asp
	365	370	375
Ser Glu Pro Val	. Leu Glu Ly:	s Asp Asp Lou Mot Asp	Met Asp Ala
	380	385	390
Ser Gln Gln Ası	i Leo Phe Asj	p Asn Lys Phe Asp Asp	Ile Phe Gly
	395	400	405
Set Ser Phe Ser	Ser Asp Pro	o Phe Asn Phe Asn Ser	Gln Asn Gly
	410	415	420
Val Asn Lys Asp	9 Glu Inya Aaդ	o His Lew Ile Glu Arg	Leu Tyr Arg
	425	430	435
Glu 1le Ser Gly	Leu Lys Ala	a Gln Leu Glu Asn Met	Lys Thr Glu
	440	445	450
Ser Gln Arg Vai	. Val Leo Gli	ı Leu Lys Gly His Val	Ser Glu Leu
	455	460	465
Glu Ala Asp Lei	Ala Glu Gla	o Gln His Lew Arg Gln	Gln Ala Ala
	470	475	480
Asp Asp Cys Gli	Phe Leu Arg	3 Ala Glu Leu Asp Glu	Leu Arg Ary
	485	490	195
Gln Arg Glu Asp	Thr Glu Ly:	s Ala Gln Arg Ser Leu	Ser Glu Ile
	500	505	510
Glu Arg Lys Ala	Gln Ala Ası	o Glu Gln Arg Tyr Ser	Гун Шеп Lys
	515	520	525
Glu Lys Tyr Son	Glw Leu Val	l Gln Asp His Ala Asp	Leu Leu Arg
	530	535	540
Lys Asn Ala Glv	Val Thm Lys	s Gln Val Ser Met Ala	Arg Gln Ala
	545	550	555
Gln Val Asp Let	Glu Arg Glu	ı Lys Lys Glu Leu Glu .	Asp Ser Leu
	560	565	570

,	WO 99.	KAORK											рстл	J\$99/11 <b>74</b> 3
			Ser	Aen	Gln	Glv	Gln	Arα	Lvs	lhr	Gln	Glu	Gln	Lev
				575					58B					585
Glu	Val	Leu	Glu	Ser 590	Ъеч	Lys	Gln	Glu	Leu 595	Ala	Tlic	Ser	Gln	Arg 600
Glu	Leu	ĞĴΠ	Val	Leu 605	Gln	Gly	Ser	Lev	610 630	Thr	Ser	Ala	Gln	5er 615
Glu	Ala	Asn	Trp	Ala 620	Ala	G1υ	Phe	Ala	G1u 625	Leu	Glu	Lys	Glu	Arg 630
Asp	Sen	Leu	Val	Ser 635	Gly	Ala	Ala	His	Arg 640	Glu	Glu	Clu	Leu	Ser 645
Ala	Leu	Arg	Lys	GI u 650	Leu	G.ln	Asp	Thr	Gln 655	Leu	Lys	Leu	Ala	Sen 660
T.)por	Glu	Glu	Ser	Met 665	Сув	Gln	Leu	Ala	Lys 670	Asp	Cln	Arg	ьуѕ	Met 675
Leu	Leu	Va1	Gly	Ser 680	Arg	Lys	Li.[A	Ala	Gใบ 685	Gln	Val.	Ile	GJ27	<b>qaA</b> 000
Ala	Leu	Asn	Cln	ьеи 695	Glu	Glu	Pro	Pro	Leu 700	Πe	Ser	Сув	Ala	Gly 705
Ser	λla	Аар	Нів	Leu 710		Ser	Thr	Val	Thr 715	Ser	rle	Ser	Ser	Cys 720
Ile	Glu	Gln	Leu	G1u 725		Ser	Тир	Ser	Gl:1 730	Tyr	Leu	Ala	Cys	Pro 735
Glu	Asp	IJe	Ser	Gly 740		Leu	l His	Ser	745		Leu	Leu	Äla	His 750
Leu	Thr	Ser	Asp	λla 755		Ala	His	Gly	AJa 760		Thr	Сув	Leu	Arg 765
Ala	Pro	Pro	Glu	Pro		. Asp	ser	Leu	775	Glu	Άla	Cys	Lys	G1n 780
Туг	Gly	Arg	; ©lu	1'hr 785		Ala	a Tyr	Leu	λla 790		Lev	G).v	Glu	Glu 795
Gly	ger	: Let	g Gla	ngA 9		ı Ası	s Sei	. Thi	· Ala 805	Met	. Arg	Asn	Cys	10 10
Şer	: Lys	Tle	e Dys	: Ala В15		e Gly	y Gl.ı	Gili	2 Let 820		Pro	Arg	Gly	Leu 825

ŕ,

WO 99/60986		-	PCT/US99/11743
_	Glu Gln≥ Leu	Gly Asp Leu Val Asp L	ys Glo Met
	830	835	840
	Ala Ala Ile	Glu Thr Ala Thr Ala A	rg IJe G1υ
	845	850	855
	Lys Ser Arg	Ala Gly Asp Thr Gly V	al Lys Leu
	860	865	870
	Arg Ile Leu	Gly Cys Cys Thr Ser L	eu Met Gln
	875	888	885
	Leo Ile Val	Ala Ser Lya Asp Leo G	Un Arg Glo
	B90	895	900
	Gly Arg Gly	Thr Ala Ser Pro Lys G	ilu Phe Tyr
	905	910	915
<del>-</del>	Arg Tup Thr	Glu Gly Leu Ile Set A	la Ser Lys
	920	925	930
_	Gly Ala Thr	Val Met Val Asp Ala A	la Asp Leu
	935	940	945
	Arg Gly Lys	Phe Glu Glu Leu Met v	al Cys Ser
	950	955	960
	Ala Ser Thr	Ala Gln Leu Va) Ala A	da Ser Lya
	965	970	975
_	Lys Asp Scr	Pro Asn Lew Ala Gln L	eu Gln Gln
	980	985	990
	Val Asn Gln	Ala Thr Ala Gly Val V	al Ala Ser
	995	1.000	1005
	Lys Ser Gln 010	lle Glu Glu Thr Asp A	sn Met Asp 1020
	Thr Leu Thr	Gin Ile Lys Arg Gln G	lu Met Asp
	025	1030	1035
	Val Leu Glu	Leu Olu Asn Glu Leu G	iln Lys Glu
	040	1045	1050
	Gly Glu Leu	Arg Lys Lys His Tyr G	In Leu Ala
	.055	1060	1065

Gly Val Ala Glu Gly Tmp Glu Glu Gly Thr Glu Ala Ser Pro Pro 1070 1075 1080

Thr Leu Gln Glu Val Val Thr Glu Lys Glu 1085 1090

- (2) INFORMATION FOR SEQ ID NO:6:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3301
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS; single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: πο
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) PEATURE: cDNA for Huntingtin-interacting protein.
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGGTGAGCTC	GAGGAGCAGC	GGAAGCAGAA	GCAGAAGGCC	CTGGTGGATA	50
ATGAGCAGCT	CCCCCACGAG	CTGGCCCAGC	TGAGGGCTGC	CCAGCTGGAG	100
CGCGAGCGGA	GCCAGGGCCT	GCGTGAGGAG	CCTCAGAGGA	AGGCCAGTGC	150
CACGGAGGCG	CGCTACAACA	AGCTGAAGGA	AAACCACAGT	GAGCTCGTCC	200
ATGTGCACGC	GGAGCTGCTC	AGAAAGAACG	CGGACACAGC	CAAGCAGCTG	250
ACCCTGACGC	$\lambda$ GCAA $\lambda$ GCCA	$\tt GGAGGAGGTG$	$\tt GCGCGGGTGA$	ACCAGCAGCT	300
GGCCTTUCAC	6'I'GGAGCAGG	TGAAGCGGGA	${\tt GTCGGAGTTG}$	AAGCTAGAGC	350
AGAAGAGCGA	CCAGCAGGAG	AAGCTCAAGA	$\tt GGGAGCTGGA$	GGCCAAGGCC	400
GGAGAGCTGG	CCCCCCCCCC	${\tt GGAGGCCCTG}$	AGCCACACAG	AGCAGAGCAA	450
GTCGGAGCTG	AGCTCACGGC	TGGACACACT	${\tt GAGTGCGGAG}$	AAGGATGCTC	500
TGAGTGGAGC	TGTGCGGCAG	CGGGAGGCAG	ACCTGCTGGC	GGCGCAGAGC	550
CTGGTGCGCG	AGACAGAGGC	GGCGCTGAGC	CGGCAGCAGC	AGCGCAGCTC	<b>6</b> 00
CCAGGAGCAG	GGCGAGTTGC	AGGGCCGGCT	$\tt GGCAGAGAGG$	GAGTCTCAGG	650
AGCAGGGGCT	GCCGCAGAGG	CTGCTGGACG	AGCAGTTCGC	AGTGTTGCGG	700
GCCCCTCCTG	CCGAGGCCGC	CEGCATCCTG	CAGGATGCCG	TGAGCAAGCT	750
GGACGACCCC	${\tt CTGCACCTGC}$	GCTGTACCAG	CTCCCCAGAC	TACCTGGTGA	800
GCAGGGCCCA	GGAGGCCTTG	GATGCCGTGA	GCACCCTGGA	GGAGGGCCAC	850
GCCCAGTACC	TGACCTCCTT	GGCAGACGCC	TCCGCCCTCG	TGGCAGCTCT	900
GACCCGCTTC	TCCCACCTGG	CTGCGGATAC	CATCATCAAT	GGCGGTGCCA	950
CCTCGCACCT	GGCTCCCACC	GACCCTGCCG	ACCGCCTCAT	AGACACCTGC	1000
AGGGAGTGCG	GGGCCCGGGC	TCTGGAGCTC	ATGGGGCAGC	TGCAGGACCA	1050
GCAGGCTCTG	CGGCACATGC	AGGCCAGCCT	GGTGCGGACA	CCCCTGCAGG	1100
GCATCCTTCA	GCTGGGCCAA	GAACTGAAAC	CCAAGAGCCT	AGATGTGCGG	1150
CAGGAGGAGC	TGGGGGCCGT	GGTCGACAAG	GAGATGGCGG	CCACATCCGC	1200
AGCCATTGAA	GATGCTGTGC	GGAGGATTGA	GGACATGATG	AACCAGGCAC	1250
GCCACGCCAG	CTCGGGGGTG	AAGCTGGAGG	TGAACGAGAG	GATCCTCAAC	1300
n'CCTGCACAG	ACCTGATGAA	GCCTATCCGG	CTCCTGGTGA	CGACATCCAC	1350
TAGCCTGCAG	AAGGAGATCG	TGGAGAGCGG	CAGGGGGGCA	GCCACGCAGC	1400
AGGAATTTTA	CGCCAAGAAC	TCGCGCTGGA	CCGAAGCCCT	CATCTCGGCC	1450
TCCAAGGCTG	TGGGCTGGGG	AGCCACACAG	CTGGTGGAGG	CAGCTGACAA	1500
GGTGGTGCTT	CACACGGGCA	AGTATGAGGA	GCTCATCGTC	TGCTCCCACG	1550
AGATCGCAGC	CAGCACGGCC	CAGCTGGTGG	CGGCCTCCAA	GGTGAAGGÜÜ	1600

	-				
AAÇAAGÇAÇA	CCCCCACCT	GAGCCCCCTG	CAGGAATGTT	CTCGCACAGT	1650
CAATGAGAGG	GCTGCCAA1G	TGGTGGCCTC	CACCAAGTCA	GGCCAGGAGC	1700
agattgagca	CAGAGACACC	ATGGATTICT	CUGGCCTGTC	CCTCATCAAG	1.750
CTGAAGAAGU	AGGAGATGGA	GACCCACCTC	$\tt CGTGTCCTGG$	AGCTGGAGAA	1800
GACGCTGGAG	GCTGAACGCA	TCCCCCTCCC	GGAGTTGCGG	AAGCAACACT	1850
ACGTGCTGGC	TGGGGCATCA	$\tt GGCAGCCCTG$	GAGAGGAGGT	GGCCATCCGG	1900
CCCAGCACTG	CCCCCCGAAG	TGTAACCACC	AAGAAACCAC	CCCTGGCCCA	1950
GAAGCCCAGC	GIGGCCCCA	GACAGGACCA	CCAGCTTGAC	AAAAACGATC	2000
GCATETACCE	AGCTCAACTC	GTGAACTACT	AGGCCCCCCA	GGGGTCCAGC	2050
AGGGTGGCTG	GTGACAGGCC	${\tt TGGGCCTCTG}$	CAACTGCCCT	GACAGGACCG	2100
AGAGGCCTTG	CCCCTCCACC	TGGTCCCCAA	GCCTCCCGCC	CCACCGTCTG	2150
GATCAATGTC	CTCAAGGCCU	$\mathtt{CTGGCCCTTA}$	CTGAGCCTGC	AGGGTCCTGG	3300
GCCATGTGGG	TGGTGCTTCT	$\tt GGATGTGAGT$	CTCTTATTTA	TCTGCAGAAG	2250
GAACTTTGGG	$\tt GTGCAGCCAG$	GACCCGGTAG	GCCTGAGCCT	CAACTCTTCA	2300
GAAAATAGTG	TTTTTAATAT	TCCTCTTCAG	AAAATAGTGT	$\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{M}\mathbf{M}\mathbf{T}\mathbf{M}\mathbf{T}\mathbf{T}$	2350
CCGAGCTAGA	GCTCTTCTTC	CTACGTTTGT	AGTCAGCACA	CTGCGAAACC	2400
GGGCCAGCGT	GGGCCTCCCT	GCCTTCTGGA	CTCCTGAAGG	TÇGTGGATGG	2450
ATGGAAGGCA	CACAGCCCGT	GCCGGCTGAT	GGGACGAGGG	TCAGGCATCC	2500
TGTCTGTGGC	CTTCTGGGGC	ACCGATTCTA	CCAGGCCCTC	CAGCTGCGTG	2550
G1'CTCCGCAG	${\tt ACCAGGCTCT}$	GTGTGGGCTA	GAGGAATGTC	GCCCATTACC	2600
TCCTCAGGCC	CTGGCCCTCG	GGCCTCCGTG	ATGGGAGCCC	CCCAGGAGGG	2700
GTCAGATGCT	GGAAGGGGCC	GCTTTCTGGG	GAGTGAGGTG	AGACATAGCG	2750
GCCCAGGCGC	TGCCTTCACT	CCTGGAGTTT	CCATTTCCAG	CTGGAATCTG	2800
CAGCCACCC	CATTTCCTGT	TTTCCATTCC	CCCGTTCTGG	CCGCGCCCCA	2850
CTGCCCACCT	GAAGGGGTGG	TTTCCAGCCC	TCCGGAGAG'I'	GGCCTTGGCC	2900
CTAGGCCCTC	${\tt CAGCTCAGCC}$	AGAAAAAGCC	CAGAAACCCA	GGTGCTGGAC	2950
CAGGGCCCTC	AGGGAGGGAC	CCTGCGGCTA	GAGTGGGCTA	GGCCCTGGCT	3000
TTGCCCGTCA	${\tt GNTTTGAACG}$	AATGTGTGTC	CCTTGAGCCC	AAGGAGAGCG	3050
GCAGGAGGGG	TGGGACCAGG	CTGGGAGGAC	AGAGCCAGCA	GCTGCCATGC	3100
CCTCCTGCTC	CCCCCACCCC	AGCCCTAGCC	CTTTMGCCTT	TCACCCTGTG	3150

CTCTGGAAAG GCTACCAAAT ACTGGCCAAG GTCAGAGGA GCAAAAAATGA 3200 GCCAGCACCA GCGCCTTGGC TTTGTGTTAG CATTTCCTCC TGAAGTGTTC 3250 TGTTGGCAAT AAAATGCACT TTCACTGTTA AAAAAAAAA AAAAAAAAA 3300

PCT/US99/11743

3301

- (2) INFORMATION FOR SEQ ID NO: 7
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 676 (B) TYPE: protein

A

WO 99/60986

- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (ix) FEATURE: Huntingtin-interacting protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
- Gly Glu Leu Glu Glu Gln Arg Lys Gln Lys Gln Jys Ala Leu Val

5 10 15

•	WO 99/60986		PCT/US99/11743
		~1	 _ 31_ 33_

'	WO 99.	/6099NO												
Asp	Asn	Glu	Gln	Len 20	Arg	Bis	Glu	Len	<b>Al</b> a 25	Gln	Leu	Arg	Ala	Ala 30
G1n	Leu	Glu	Arg	<b>G</b> lu 35	Arg	Ser	Gln	Gly	Leu 40	Arg	Glu	Glu	Ala	Glu 45
Arg	Lys	Ala	Ser	Ala 50	ፐስተ	Glu	Ala	Arg	Tyr 55	Asn	Lys	L∉u	Lys	Glu 60
TÀS	His	Ser	Glu	Leu 65	Vəl	Hip	Və ).	Big	Ala 70	Glu	Lev	Leu	Arg	Lув 75
Asn	Ala	Aar	Thr	Ala 80	Lys	Gln	Leu	Thr	val 85	Thr	Cln	alD	Ser	Gln 90
Glu	Glu	۷al	λla	Arg 95	Val.	ГАВ	Glu	αlĐ	Lena 100	Ala	Phe	Glm	Va.l	Glu 105
Gln	Val	Ľys	Arg	Glu 110	Ser	Glu	Leu	Ьув	Leu 115	Glu	Glu	Lys	Ser	Авр 120
Gln	Gln	Glu	Lys	Leu 125	Lys	Ary	Glu	Len	Glu 130	Ala	Lуs	Ala	Gly	Glu 135
Legn2	Ala	Arg	Ala	Gln 140	G3 u	λla	Leu	Ser	His 145	Thx	Glu	σ£ο	Ser	Љув 150
Ser	Glra	Leu	Ser	Ser 155	Arg	Leu	Asp	Thr	ьеи 160	Ser	λla	Glu	Lys	Авр 165
Ala	Leu	Ser	Gly	Ala 170	Væl.	Ang	Gln	Arg	Glu 175	Ala	Asp	Leu	Leu	Ala 180
Ala	Glπ	Sex	. Tën	Val 185	Arg	Glu	Thr	Glu	Ala 190	Ala	Ъeu	Ser	Ärg	Glu 195
Gln	Gln	Arg	Ser	Sex 200	Glm	Glu	Gln	Gly	Glu 205	Leu	Gln	Gly	Arg	Lou 210
Ala	. Glu	Arg	r Glu	Ser 215		Glu	(Gin	Gly	Նeա 22D		Gln	Axg	Leu	Leu 225
Asp	Glu	Gln	n Phe	ala 230		Leu	ı yığ	Gly	Ala 235		Ala	Glu	Ala	Ala 240
Gly	, 11e	: Lev	Gl.n	Авр 245		Va)	L Ser	Lys	Leu 250		Asp	Pro	ь <del>е</del> и	His 255
Leu	arg	Cys	Thr	Ser 260		Pro	дай с	Туп	1.eu 265		. Ser	Ar <u>c</u>	Ala	Gln 270

•	WO 99	/60986								<u> </u>			PCIN	US99/11743
Glu	Ala	Leu	Asp	Ala 275	Val	Ser	Thr	Leu	Glu 288	Glu	ĠĴλ	His	Ala	Gln 285
Tyr	Leu	Thir	Ser	Leu 290	Ala	Asp	Ala	ser,	Ala 295	Leu	Val	Bla	Ala	Leu 300
Thr	Arg	Phe	Sex	Н18 305	Len	Ala	Alæ	Asp	70 310	Ile	Ile	Азп	Gly	Gly 315
Ala	Thr	Ser	His	Leu 320	Ala	Pro	Thr	Asp	Pro 325	Ala	Asp	Arg	Leu	Ile 330
Asp	Thr	Сув	Arg	Glu 335	Cys	Gly	Ala	Arg	Ala 340	Leu	Glu	Leu	Met	Gly 345
Gln	Leu	Gln	Asp	Gln 350	Gln	λla	Leu	Arg	His 355	Met	Gln	Ala	Ser	Leu 360
Va1	Arg	Thr.	Pro	<b>Бе</b> ш 365	Gln	Gly	Ile	Leu	Gln 370	Leu	GJA	Gln	Glu	Leu 375
Lys	Pro	Lys	Ser	leu 380	Asp	Va l.	Arg	Gln	G1u 385	Glu	Leu	Gly	Ala	Val 390
Val	Asp	Lys	Glu	Met 395	λla	λla	Thr	Ser	<b>A</b> la 400	Ala	Ile	Glu	Asp	Ala 405
Val	Arg	Arg	Il.e	Glu 410	Авр	Met.	Met.	Asn	Gln 415	Ale	Arg	His	Ala	Ser 420
Ser	GJA	Val	Lys	ьеи 425	Glu	Val	Asn	Glu	Arg 430	Île	Leu	Asn	Ser	Сув 435
Thr	ABD	Leu	Met	Lys 440	Ala	Ile	Arg	Leu	<b>L</b> eu <b>44</b> 5	Val	Tłìr	Thr	Ser	Thr 450
Ser	ren	Gln	ьуѕ	G1 ນ 455	Ile	Val	Glu	Ser	Gly 460	Arg	Gly	Mla	A1a	Thr 465
Gln	G].n	Glu	Phe	Туг 470	Ala	Lув	Aen	Ser	Arg 475	Ттр	Thr.	Glu	Gly	Leu 480
Ile	Ser	Ala	Ser	<b>Ъув</b> 485	Ala	Val	Gly	Trp	Gly 490	Ala	Thr	Gln	Leu	Va ). 495
Glu	Ala	Ala	Авр	1.ya	Val.	Val	Lena	Hi.s	Thr: 505	G].Y	Гув	Tyr	Glu	Glu 510
Leu	Ile	Val	ርንደ	ser 515	His	Glu	Ilc	Ala	Ala 520	Ser	Thr	Ala	Gln	Leນ 525

					)									
•	NO 99/	60986											PCT/U	899/11743
Va ĵ.	Ala	Aìa	Sex	Lys 530	Val	Lув	Ala	Asn	<b>Бу</b> в 535	Нів	Ser	Pro	ні.в	Len: 540
Ser	Arg	Leu	Gln	G1u 545	Сув	Ser	Arg	ግጥr	Val 550	Asn	Glu	Arg	Ala	Ala 555
Aso	Val	Val	Ala	Ser 560	Thr	ρλε	Ser	сіу	Gln 565	Glu	Gln	Ilc	Glu	Asp 570
Arg	Asp	Тінг.	Met	Asp 575	Phe	Şer	ĠĵĀ	Leu	Ser 588	Leu	I.le	Lys	Leu	Lys 585
Lys	Gln	Glu	Met	Glu 590	Thr	Gln	Val	Arg	<b>v</b> al 595	Leu	Clu	Lou	Glu	Lys 600
Thr	Len	Glu	Ala	Glu 605	Arg	Met	Ary	Len	610 GjA	Glu	Læu	Ang	lys	Gln 615
His	Туг	Val.	Leu	Ala 620	Glγ	Ala	Ser	Gly	Ser 625	Pro	Gly	Glu	Glu	Va1 630
Ala	IJë	Ang	Pro	ser 635	Tłźr	Alu	Pro	Arg	Ser: 640	ya).	Thir	Thu	Lys	Бун 645
Pro	Pro	Leu	A.J.a	G l n 650	lys	Pro	Ser	Val.	Ala 655	Pro	Arg	Gla	qaK	Нів 660
Gla	Leu	Asp	ьуз	ьув 665		GJY	ıle	Ί'nχ	Pro 670	Ala	Gln	Leu	Val	λ <del>s</del> n 675
Туг														
			ON F				3:							

(i) \$EQUENCE CHARACTERISTICS:

(A) LENGTH: 2338

(B) TYPE: nucleic acid.

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: mouse

(ix) FBATURE: cDNA for Huntingtin-interacting protein - mHIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGCACGAGGG CTCATTCAGA TCCCCCAGCT GCCCGAGAAT CCACCCAACTT 50 CCTACGAGCC TCGGCCCTGT CAGAGCACAT CAGTCCTGTG GTGGTGATCCC 100 GGCAGAGGTG TCATCCCCAG ACAGTGAGCC TGTCCTGGAG AAGGATGACCT 150 CATGGACATG GACGCCTCCC AGCAGACTTT GTTTGACAAC AAGTTTGATGA 200

WO 99/6098	6			PCT	/US99/11743
CGTCTTTGGC	AGCTCATTCA	GCAGCGACCC	TTTCAATTTC	AACAATCAAAA	250
TGGCGTGAAC	AAGGACGAGA	AGGACCACTT	GATTGAACGC	CTGTACAGAGA	300
GATCAGTGGA	CTGACAGGCC	AGCTGGACAA	CATGAAGATT	GAGAGCCAGCC	350
GCCATGCTG	CAGCTGAAGG	GTCGAGTGAG	TGAGCTGGAG	GCAGAGCTAGO	400
AGAGCAGCAG	CACTTGGGCC	GGCAGGCTAT	GGATGACTGC	GAGTTCCTGCG	450
CACTGAGCTG	GATGAACTGA	AGAGGCAGCG	AGAGGACACG	GAGAAGGCACA	500
GCGCAGCCTG	ACTGAGATAG	AAAGAAAGGC	CCAGGCTAAT	GAACAGAGGTY	550
TAGCAAGTTA	AAAGAGAAGT	ACAGTGAACT	GGTCCACAAC	CATCCTGACCT	. 600
GCTGCGGAAG	AACGCAGAGG	TGACCAAACA	GETETECETE	GCCCGCCAACC	650
CCAGGTGGAT	TTGGAAAGAG	AGAAAAAAGA	GUTAGCAGAT	TCCTTTGCAC	700
GTGTAAGTGA	CCAGGCCCAG	$\tt CGGAAGACTC$	AAGAGCAACA	GGATGTTCTA	750
GAGAACCTGA	AGCATGAACT	$\tt GGCCACCAGC$	AGACAGGAGC	TGCAGGTCCT	800
CCACAGCAAC	CTGGAAACCT	$\tt CTGCCCAGTC$	AGAAGCGAAA	TGGCTGACAC	850
AGATCCCCGA	GTTGGAGAAG	GAACAAGGCA	GCTTGGCGAC	TGTTGCAGCT	900
CAGAGAGACC	AAGAGTTATC	AGCCCTCCGA	GACCAGCTGG	AAAGCACCCA	950
GATCAAGCTG	CCTCGCCGCCC	AGGAATCCAT	GTGCCAGCAG	CTGAACGACC	1000
AGAGGAAAAC	CCTCTTGGCA	CCCATCAGGA	ACCCTGCGGA	GCGTGAGATA	1050
CAGGAGGCGC	TGAGCCAGCT	TGAGGAACCC	${\tt ACCCTCATCA}$	GCTGTGCAGG	1100
ATCCACAGNT	CACCTTCTCT	CCAAAGTCAG	$\mathtt{CTCCGTTTCC}$	AGCTGCCTCG	1150
<b>AGCAAUT</b> GGA	MANGAACGGC	AGCCAGTATC	TGGCCTGCCC	AGAAGATATT	1200
AGTGAGCTTV	TGCACTCGAT	CACCCTGCTT	GCCCACTTGA	CCGGTGACAC	1250
TGTCATCCAG	GGGAGTGCCA	CCAGCCTCCG	GGCCCCACCG	GAGCCAGCCG	0.300
ACTCGTTGAC	CGACGCCTGT	AGGCAGTATG	GCAGAGAAAAC	CCTGGCCTAT	1.350
CTGTCCTCCC	TGGAGGAAGA	GGGAACTGTG	GAGAATGCTG	ACGTCACAGC	1400
CCTTAGGAAT	TGCCTCAGCA	CCCTCAAGAC	CCTTGGCGAG	GAGCTGCTGC	1450
CCAGGGGCCT	GGACATCAAG	CAGGAAGAGC	${\tt TGGGTGACCT}$	GGTGGACAAC	1500
GAGATGGCAG	CCACTTCAGC	TGCCATTGAA	GCTGCCACCA	CCCGGATAGA	1550
GGAAATTCTC	AGTAAGTCCC	GAGCAGGAGA	CACGGGAGTC	AAGCTGGAGG	1600
TGAATGAGAG	GATCCTGGGT	TCCTGTACCA	GCCTGATGCA	GGCCATCAAG	1650
GTGCTCGTTG	TGGCCTCCAA		ANGGNGATAG	TGGAGAGTGG	1700
CAGGGGTAGT	GCATCCCCTA		CGCCVVGVVC	TCTCGGTGGA	1750
CGGAAGGGCT	GATATCCGCC	TCCAAACCTC	TYCGTTGGGG	AGCTACCATC	1800
		TGTGGTCCAA			1850
				CAGCTCCTCC	
				GACCCAGCTG	
				TGGTGGCCTC	
				ATGGACTTCT	
				TTCCCAGGTT	
				AGAAACTAGG	
AGAGCTACGG	AAGAAACACT	ACGAGCTGGA	GGGCGTGGCT	GAGGGCTGGG	2250

AGGAAGGGAC AGAAGCATCA CCGTCTACTG TCCAAGAAGC AATACCGGAC 2300

2338

## (2) INFORMATION FOR SEQ ID NO; 9:

AAAGAGTAGA GCCAAGCCGA CACCCCACAC ATCAGAAA

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 676
- (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein

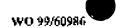
- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: mouse
- (ix) PEATURE: Huntingtin-interacting protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
- Ala Arg Gly Leu Ile Gln Ile Pro Gln Leu Pro Glu Asn Pro Pro 5 10 15
- Asn Phe Leu Ang Ala Sem Ala Leu Sem Glu His Ile Sem Pro Val 20 25 30
- Val Val lle Pro Ala Glu Val Ser Ser Pro Asp Ser Glu Pro Val 35 40 45
- Leu Glu Lys Asp Asp Neu Met Asp Met Asp Ala Ser Gln Gln Thr
  50 55 60
- Lou Phe Asp Asn Lys Phe Asp Asp Val Phe Gly Ser Ser Leu Ser 65 70 75
- Ser Asp Pro Phe Asm Phe Asm Asm Glm Asm Gly Val Asm Lys Asp 80 85 90
- Glu Lys Asp his Lew Ile Glu Arg Lew Tyr Arg Glo Ile Ser Gly 95 100 105
- Leu Thr Gly Gln Leu Asp Asn Met Lys 11e Glu Ser Gln Arg Ala 110 115 120
- Mot Leu Gln Leu Lys Gly Arg Val Ser Glu Leu Glu Ala Glu Leu 125 130 135
- Ala Glu Glu Glu Ris Leu Gly Arg Glu Ala Met Asp Asp Cys Glu 140 145 150
- Phe Leu Arg Thr Glo Leu Asp Glu Leu Lys Arg Glu Arg Glu Asp 155 160 165
- Thr Glu Lys Ala Gln Arg Scr Leu Thr Glu ile Glu Arg Lys Ala 170 175 180
- Gln Ala Asn Glu Gln Arg Tyr Ser Lys Leu Lys Glu Lys Tyr Ser 185 190 195
- Glu Leo Val Gln Asn His Ala Asp Leo Leo Arg Lys Asn Ala Glu 200 205 210
- Val Thr Lys Gln Val Ser Val Ala Arg Gln Ala Gln Val Asp Leu 215 220 225
- Glu Arg Glu Lys Lys Glu Leo Ala Asp Ser Phe Ala Arg Val Ser

	WO 99	/60986											PCT/	U <b>8</b> 99/1:
				230					235					240
Asp	Gln	λla	Gln	Arg 245	Lys	Thr	Gln	Glu	Gln 250	Gln	Анр	Va l	Leo	Glo 255
Asn	J,eu	Lys	His	Glu 260	Leu	Ala	ТЭ1т	Ser	Λrg 265	Glis	Glu	ьеи	Gln	Val 270
Leu	His	Ser	Asn	Leu 275	Glu	Thr	Ser	Ala	Gln 288	Ser	Glu	Ala	Lys	Trp 285
Leu	Thr	Gln	Ilc	<b>A</b> la 290	Glu	Геп	Сĵп	Lуя	Glu 295	Gln	GTY	Ser	Leu	Ala 300
Thr	Və)	Ala	Ala	G <b>l</b> n 305	Arg	Gliu	<b></b> ՅՂԱ	Glu	Leu 310	Ser	Ala	ьeu	Arg	Asp 315
Gln	Leu	Glu	Ser	195r 320	Gln	Ile	Lys	Len	Ala 325	Glγ	Ala	Gln	Glu	Ser 330
Met	Cys	Glu	Gln	Val 335	Lys	qa <b>A</b>	Gln	улд	Lys 340	Thr	Leu	Leu	Ala	Gly 345
Ile	Arg	Lys	λla	Ala 350	Glu	Arg	Glu	lle	Gln 355	Glu	Ala	Leu	Ser	Gln 360
Leu	G.l 13	Glu	Ртю	Thr 365	Leu	Ile	Ser	Cys	Ala 370	GJA	Ser	<b>T</b> h-	ABD	ні в 375
Leu	ьеи	Ser	Ьуз	Val 380	Ser	Ser	Val	Ser	Ser 3B5	Cys	Leu	Glu	Gln	Leu 390
Glu	Lys	Asn	Cly	Ser 395	©ln	ТУΥ	Leu	Ala	Сув 400	Pro	Glu	Asp	lle	Ser 405
Glu	Leu	Leu	His	Ser 410	Ile	Thr	Leu	Leu	λla 415	Kis	Len	Thr	Gly	Asp 420
Thr	Val	Ile	Glv	Gly 425	Ser	Ala	Thr	Ser	Len 430	Arg	Ala	Ρτο	Pro	Gl ₁₂ 435
Pro	Ala	Asp	Ser	Leu 440	<b>ፓ</b> ስነ-	Glo	A1a	Суз	Arg 445	Gln	Tyr	GТУ	λrg	Glu 450
Thr	Leu	Ala	тут,	Leu 455	Ser	Ser	Leu	Glu	Glu 460	Glız	Ġĵλ	Tłik	Val.	Glu 465
Asn	Ala	qaA	Val	Thr 470	Ala	Leu	Arg	Asn	Cys <b>4</b> 75	Len	Ser	Arg	Val	Lys 480

	NO 99.	ie nobe			)								PCT/U	899/11743
			Glu	Glu 485	Leu	Leu	Pro	Arg	Gly 490	Leu .	Asp	lle	ГЛа	Gln 495
Glu	Glu	Leu	Cly	Asp 500	Leu	Va1	Авр	Lys	G.Iv 505	Met	Ala	Al.a	Thx.	Ser 510
Ala	Ala	Ile	Glu	Ala 515	Λla	בנלים-	That	Arg	Il.e <b>52</b> 0	Glv	Glu	Ile	I'ëii	Ser 525
Lys	Ser	Arg	Ala	G3. <b>y</b> 530	Asp	Thr	Gly	Val	Lys 535	Leu	Clu	Val	Asn	Glu 540
Arg	Ile	Leu	Gly	Ser 545	Сув	Thr	Ser	Leu	Met. 550	G) n	Ala	T.) e	lys	Val 555
Leu	Val	Va I	Ala	Ser 560	Lys	Asp	Leu	Gln	ьуs 565	Glu	11c	Val	Glu	Ser 570
Gly	Arg	Gly	Ser	Alla 575	Ser	Pro	Lys	Glu	Phe 588	Tyr	Ala	Lys	Asn	Ser 585
Arg	Ттр	Thr	Glu	<b>G</b> Ј.у 590	Leu	Ilc	Ser	Ala	Ser 595	Lys	Ala	Va1	Gly	Trp 600
Gly	Ala	Thr	Ile	Met 605	Val.	Анр	Ala	Ala	Asp 610	Leu	<b>v</b> al	Val	Cln	Gly 615
Lув	Gly	Lys	Phe	61u 620	Glu	Leu	Met	Val	Сув 625	Ser	Ar'g	Glu	IÌë	<b>Ala</b> 630
Ale	Ser	uhr:	: Ala	Gln 635	Len	Val	Ala	Ala	Ser 640	ГЛЗ	Val	Ъус	λla	<b>Դ</b> Ֆր 645
Lγε	Gly	/ Sei	. Leu	asa 650	Leu	מתוי	Gln	Leu	GJ.n 655	Glu	A Ì.ia	Ser	Arg	Gly 660
Val	. Авг	ı Gl.ı	ı Ala	Thr 665		Ala	Val	Val	Λla 670	Ser	Thr	Ile	Sen	G) y 675
Lys	క న్లు	r Gli	n Ile	680		Thr	Asp	Ser	Met 685	day :	Phe	Ser	Ser	Met 690
Tìn	c Let	ı Thi	r Glr	11e 695		arg	Gln	Glu	. <b>Me1</b> 700	. Asp	Ser	Glm	Val	Arg 705
Va:	l Lei	o Gl	u L≌v	710		a Asp	Len	Glr	. <b>Lys</b> 715	; Glu	Arg	Gln	і І.ув	A50 Pea
Gl	y Gl	u Le	u Arq	у <b>Бу</b> я 725		His	з Туі	Gli	2 Lev 73(	ı Glu	Gly	Val	. Ala	Glu 735

1.450

1500



Gly Trp Glo Glo Gly Thr Glo Ala Ser Pro Ser Thr Val Glo Glo 740 745

Ala Ile Pro Asp Lys Glu 755

- (2) INFORMATION FOR SEQ ID NO:10:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3964
- (B) TYPE: nucleic acid.
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM; mouse
- (ix) FRATURE: cDNA for Huntingtin-interacting protein mHIP1a
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: [0:
- GGCACGAGGC GGCGCGCGC CTCCGTGTGC CTAGGCTTGA GGCGGGGGGCGT 50 DTACCACO OCCUCACIÓN OSCOCOSO SASSOCIOTO TACTOCOACA 100 AACAGCATCA AGAATGTGCC GGCGCGGGTC CTGAGCCGCA GGCCGGGCCA 150 CAGCCTAGAG GCCGAGCGCG AGCACTTCGA CAAGACGCAG GCCATCAGTA 200 TCAGCAAAGC CATCAACAGC CARRAGGCC CAGTGAAGGA GAAGGATGCC 250 CGGCGTATCA TCCTGGGCAC GCATCATGAG AAGGGAGCCT TCACCTTCTG 300 GTCCTATGCC ATCGGCCTGC CGCTGTCCAG CAGCTCCATC CTCAGCTGGA 350 AGTTCTGTCA CGTCCTTCAC AAGGTCCTCC GGGACGGACA CCCCAACGTC 400 CTGCATGACT ATCAGCGGTA CCGGAGCAAC ATACCTGACA TCGGTGACTT 450 GTGGGGCCAC CTTCGTGACC AGTATGGACA CCTCGTCAAT ATCTATACCA 500 AACTGTTGCT GACTAAGATC TCCTTCCACC TTAAGCACCC CCAGTTTCCT 550 GCAGGCCTGG AGGTAACACA TGACCTGTTG GAGAAGGGGG CGGGAACTGA 600 TGTYAACAAC ATTTTTCAGC TTACCGTGGA GATGTTTGAC TACATGGACT 650 GTGAACTGAA GCTTTCTGAG TCAGTTTTCC GGCAGCTCAA CACGGCCATC 700 GCAGTGTCCC AGATGTCTTC TGGCCAG/2G/2 CCCCTAGCGC CCCTCATCCA 750 GGTCATTCAG GACTGCAGCC ACCTCTACCA CTACACAGTG AAGCTCATGT 800 TTAAGCIGCA CICCTGTCTC CCGGCAGACA CCCTGCAAGG CCACAGGGAT 850 CGGTTCCACC ACCAGTTCCA CAGCCTCAAA AACTTCTTCC GCCGGGCTTC 900 AGACATGCTG TACTTCAAGA GGCTCATCCA GATCCCGCGG CTGCCTGAGG 950 GACCCCCAA TTTCCTGCGG GCTTCAGCCC TGGCTGAGCA CATCAAGCCG 1000 GTGGTGGTGA TTCCCGAGGA GGCCCCAGAG GAAGAGGAGC C'IGAGAACC'I' 1050 AATTGAAATC AGCAGTGCGC CCCCTGCTGC GGACCCACTG CTCCTGCCTG 1100 ACCTCTTGA TCAGACCTTT GGACCCCCCA ATGGCTCCAT GAAGGATGAC 1.1.50 AGGGACCTCC AAATCGAGAA CTTGAAGAGA GAGGTGGAGA CCCTCCGTGC 1200 TGAGCTGGAG AAGATTAAGA TGGAGGCACA GCCC1'ACA'1C TCCCAGC'1GA 1250 AGGGCCAGGT GAATGGCCTG GAGGCAGAGC TGGAGGAGCA GCCCAACCAC 1300 AAGCAGAAGG CCCTGGTGCA LAALGAGCAG CTGCGCCACG AGCTGGCCCA 1350 GCTCAACCC CTCCAGCTG AGGGCGCCG CAACCAGGC CTTCGAGAGG 1400 AAGCAGAGAG GAAGGCCAGT GCCACGGAGG CACGCTACAG CAAGCTGAAG

GAGAAACACA GCGAACTCAT TAACACGCAC GCCGAGCTGC TCAGGAAGAA

WO 99/60986 PCT/US99/11743

ALC: Williams					
CGCAGACACG G				01.0000	1550
TGGCACGGC'l' A			AGATGGAGCA	<b></b>	1600
GAGTOTGAGA TY					1650
GAGGGAGCTG G					1700
TGAGCCGCAC A					1750
CTGAACGCGG A		CCTCACTCA			1800
AGAGCTGCTG G					1850
GCCANGAGCA G	~	TCCCAGGAGA			1900
CTGGCAGAAA A	GGAGTCTCA	GGAGCAGGGG	CTTCGGCAGA	,., _ 0 0 .,	1950
TGAGCAGTIG G	CGGTGTTGC		CGCCGAGGCA	GAGGCCATCC	2000
TACAGGATGC A	GTGAGCAAG		CCCTGCACCT	CCGCTGCACC	2050
AGCTCCCCAG A			CAGGCAGCCC		2100
GAGCGGCCTG G	AGCAGGGCC	ACACCCAGTA	CCTGGCTTCC	TCCGAAGATG	2150
CTTCTGCCCT G	GTGGCAGCG	CTGACCCGCT	TCTCCCATTT	GGCTGCGGAC	2200
ACCAPTIGTCA A	TGGTGCCGC	${\tt CACCTCCCAC}$	ADDDDDDDTD	CCGACCCCGC	2250
CGACCGCCTG A	TGGACACAT	GCAGGGAGTG	TGGAGCCCGG	GCTCTGGAGC	2300
TGGTGGGACA G	CTGCAAGAC	${\tt CAGACAGTGC}$	TACGGAGGGC	TCAGCCCAGC	2350
CTGATGCGCG C	ADDTODOO:	GGGCATTCTG	CAGTTGGGCC	AGGACTTGAA	2400
GCCTAAGAGC C	TEGATETAC	GGCAAGAGGA	GCTAGGGGCC	ATGGTGGACA	2450
AGGAGATGGC G	GCCACCTCG	GCAGCCATTG	AGGACGCTCT	CCCGAGGATC	2500
GAGGACATGA T	GAGCCAGGC	CCGCCACGAG	AGCTCAGGCG	TGAAACTGCA	2550
GGTGAATGAG A	GGATCCTCA	ACTCCTGCAC	AGACCTGATG	AAGGCTATCC	260D
GGCTCCTGGT G	SATGACCTCC	ACCAGCCTGC	AGAAGGAAAT	TGTGGAGAGC	2650
GGCAGGGGGC C	AGCANCGCA	GCAGGAATTT	TATGCCAAGA	ATTCACGGTG	2700
	PATICICAG	CCTCTAAGGC	AGTGGGCTGG	GGAGCCACAC	2750
AGCTGGTGGA G	TCACCTGAC	AAGGTTGTGC	TTCACATGGG	CAAATACGAG	2800
		TGAGATTGCG		CCCAGCTGGT	2850
GGCAGCCTCG A	<b>DAAADTODA</b> A	ССАЛСАЛДАЛ	CAGTECECAC	TTGAGCCGCC	2900
TGCAGGAATG T	TOCOGGRACT	GTCAACGAGA	GGGCTGCCAA	CGTCGTGGCC	2950
TCCACCAAAT C	TGGCCAGGA	GCAGATTGAG	GACAGAGACA	CCATGGATTT	3000
CTCTGGCUTC T	PCCCTCATCA	AGTTGAAGAA	GCAGGAGATG	GAGACACAGC	3050
	KFAGCTGGAG	AAGACACTAG	AGGCAGAGCG	TGTCCGGCTC	3100
GGGGAGCTTC G	gaaacagca	CTATGTACTG	GCTGGGGGGA	TGGGAACACC	3150
TAGCGAAGAA	ADDADOSAE	GACCCAGCCC	AGCTCCCCGA	agtggggcca	3300
CTANGAAGCC A	ACCGCTGGCC	CAGAAACCCA	CCATACCCCC	CAGGACAGAC	3250
AACCAGCTCGA	CAAAAAGGA	T GGTGTCTAC	C CAGCTCAAC	OMTOAADTOT T	3300
TAGGCCCCTAA			G GTGGTTGTG	C CTGGGGCTTCA	3350
TGTGGCTGTCT	GGCAGTGGT	C AAGGGGCCT	C TGAGAAGCC	T CCAACTCCTG	3400
CCCAAGGGGCC	TACTCTGTG	G GACAGTTCA	T CTGGATGTG	A ATCTATTAT	3450
CTTAAGTAGGA		A CCAGCTGGG	A CCCAGCAGG	C CTGAGCCACA	3500
AATCTGCAGCG			A TOUTGUGAG	G TATTICITTC	3550
TTÇGTAAGTTI			G GTCACATAA	G CCAGGAGCCT	3600
CCTTGTCTCTG			O <mark>AAGTGAA</mark> T T!	A ACAGAAAGAG	3650
GGTCCCTGCTG			T GACCTGTGA	C CCTTGAGCCA	3700
GGGAGAGCAGG	-		A CCTGGGGGC	C TGGTGCTAGG	3750
GCATCCATGCT	GGGAGCCCC	A CCAGACCAG	G CTTTGTGTGTC	C GAGCCTGGCA	3800
TCATCGTGGCT	GGGGCAGCC	C UTGCTCAGO	T COTOTOTO	CCCGTGACCT	3850
TGAAGCCACCC		A CAGITITCO	A TTCTCCTGG	C TACTAGTGTG	3900
ርሮፕሮኖጥንልጥግር	CCTACCTTG	A TGAGTAGAT	TOAGCCCTC	DDDTTDDAAAT D	3950
GCCTTTCCTCG		<b></b>	<del>_</del>		3964
20011100100					

- (2) INFORMATION FOR SBQ ID NO: 11:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTII: 676
- (B) TYPE: protein
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: mouse
- (ix) FEATURE: Huntingtin-interacting protein -mHIP1a
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met	λεп	Ser	Ile	Lys 5	Asn	Va1	Pro	λla	Arg 10	Val	Ľеи	Ser	Arg	Arg 15
Ρπο	ĞЈУ	His	Ser	Leu 20	Glu	Ala	Ģlu	Ang	G] w 25	G] n	Phe	qsA	Iys	Thr 30
Gln	A l.a	I.l te	Ser	11∉ 35	Ser	Lys	Ala	Лe	Asn 40	Ser	Glη	Gkυ	Ala	Pro 45
Val	Lys	Glu	гЛа	nis 50	λla	Arg	Arg	lle	Tle 55	Leu	сзу	TÛDY.	His	Bis 60
Glu	Lys	Gly	A l.a	Phe 65	Tža <del>n</del>	Phe	TT)	Ser	ጥ <del>ታ፣</del> 70	A l.a	I).e	СĴУ	Len	Рто 75
Leu	Ser	Ser	Ser	Ser BO	lle	Leu	Ser	Trp	ь <b>у</b> в 85	Phe	Cys	His	Val	Leu 90
Hie	Lys	Val	Leu	Arg 95	Asp	Gly	His	Pro	Asn 100	Val	Len	His	Asp	Туг 105
Gln	Arg	Tyr	Arg	Ser 110	Asn	ılc	Arg	Clu	11e 115	Gly	Asp	Leu	TTP	Cly 120
His	Leu	Ary	Agp	Gln 125	Туг	Gly	Вів	Len	Val 130	Asn	Ile	Тут,	Thr	Lys 135
Leu	Leu	Leu	Thr	Lys 140	Ile	Ser	Phe	Hís	Leu 145	Lys	His	Pro	Gln	Phe 150
Pro	Ala	Gly	Leu	Glu 155	Val	The	Asp	G) u	Val. 160	3.e11	G) 11	Lуь	Alla	Ala 165
Gly	Τ'nι	Asy	Val	Asn 170	Asn	Ile	Phe	Gln	Leu 175	Thr	<b>V</b> al	Clu	Mct	Phe 180
Asp	туг	Met	Asp	Сув	Glu	Len	Jyr	Len	Ser	Glu	Ser	Val	Phe	Arg

WO 99/60986		

WO 99/60986				PCT/US99/11743
170 33100300	185	190		195
Glm Leu Asn Thr	Ala Ile Ala 200	Val Ser Gln 205	Met. Ser Ser	Gly Gln 210
Cys Ary Leu Ala	Pro Leu Ile 215	Gln Val Ile 220	б∣п Авр Сув	Ser His 225
Leu Tyr His Tyr	Thr Val Tys 230	Neu Met Phe 235	Lys Leu Ilis	Ser Cys 240
Leu Pro Ala Asp	Thr Leu Gln 245	Gly His Arg 250	Asp Arg Phe	Нів Glю 255
Glm Phe His Ser	Leo Lys Agn 260	Phe Phe Arg 265	λrg λla Ser	Asp Met 270
Leu Tyr Phc Lys	Arg Leu ile 275	Gln 1le Pro 288	Arg Leu Pro	Glu Gly 285
Pro Pro Asn Phe	Leu Arg Ala 290	Ser Ala Leu 295	Ala Glu His	Ile Lys 300
Pro Val Val Val	ile Pro Glu 305	Glu Ala Pro 310	Clu Glu Glu	Clu Pro 315
Glu Asn Leu Ilc	Glu 1le Ser 320	Ser Ala Pro 325	Pro Ala Cly	Clu Pro 330
Val Val λla	Asp Leu Phe 335	Asp Gln Thx 340	Phe Gly Pro	Рто Asn 345
Gly Ser Met Lys	Asp Asp Arg 350	Asp Leu Gln 355	Ile Glu Asn	Leu Lys 360
Arg Glu Val Glo	7hr Len Arg 365	Ala Glo Leu 370	Glu Lys Ile	Lys Met 375
Glu Ala Gln Arg	Tyr Tle Ser 380	Gln Leu Lys 385	Cly Cln Val	Asn Gly 390
Lou Glu Ala Glu	Leu Glu Glu 395	Gln Arg Lys 400	Gln Lys Gln	Lys Ala 405
Lev Val Asp Asn	Glu Cln Leu 410	Arg His Glu 415	Leu Ala Cln	Leu Lys 420
Ala Leu Gln Leu	Glu Gly Ala 425	Ang Asn Gln 430	Gly Leu Arg	Glo Glu 435
Ala Glu Arg Lys	Ala Ser Ala	Thr Glu Ala	Arg Tyr Ser	Lys Leu

	,	
WO	<b>QOARHURK</b>	

1	WO 99/	00986								_			PCIR	ווזקקטן
				440					445					450
Lys	Glu	Lys	нія	Ser 455	Glu	Leu	IJe	Asn	Thr 460	Ні. в	Ala	Glu	ьeu	ьеи 465
Arg	Lys	Asn	Ala	Лэр 470	Thr	Λla	рÀг	Gln	Leu 475	Thr	Val	Thr	Gln	Gln 480
Ser	G.lrs	Glu	Glu	Val 485	Ala	gıA	Val	Lys	Glu 490	G.l.n	Lėji	Ala	Phe	Gln 495
Met	Gln	GJπ	a.í A	Lун 500	Arg	Glu	Ser	Glu	Met. 505	Lys	Met.	Glu	Glu	Gln 510
Ser`	Asp	Gln	ьсп	Glu 515	Ľys	Leu	Lys	Arg	Glu 520	Leu	Ala	ts [A	Arij	Ala 525
Gly	Glu	Leu	Ala	Arg 530	Ala	ូβ].π	G)ນ	Ala	Leu 535	Ser	Arg	Thr	Glu	Gln 540
Ser	Cly	Ser	Glu	Leu 545	Ser	Ser	yrg	Leu	Asp 550	ገስድ	bou	Asn	Ala	Glu 555
Гув	Gju	Ala	Leu	Ser 560	Gly	Val	Val	Ату	Gln 565	Arg	Glu	Ala	G).u	1.eu 570
Leu	Ala	sīA	Gln	Ser 575	Leu	Val	Arg	Glu	<b>Lys</b> 588	Glu	Glu	λla	Leu	Ser 585
Gla	Glu	Glu	Gln	Arg 590	Ser	Ser	Cln	Clu	Lys 595	CJA	Clu	Leu	guá	01y 600
Gln	Leu	λla	Glo	1₁уя 605	Glu	Ser	Gln	Glu	Gln 610	Gly	Leu	Ат:ц	Gla	I.ys 615
Leu	Len	Asp	Glu	Gln 620	Leu	Ala	val	L@u	Arg 625	Ser	Ala	Ala	Alä	Glu 630
Ala	Glu	λla	Ile	Lev 635	G,).n	qaA	Ala	Val	Ser 640	Lys	Leu	Asp	Asp	Pro 645
Leu	His	Leu	Arg	Cys 650		Ser	Ser	Pro	Asp 655	Туr	Leu	Val	Ser	Arg 660
Ala	Cln	Ala	Ala	ьеи 6 <b>6</b> 5		Ser	Val	Ser	Gly 670	Leu	Glu	Gln	Gly	His 675
Thr	G1n	ፓሃ፣	Lev	<b>A</b> )a 680		Ser	Glu	qaA	Ala 685	Ser	Ala	Leu	Va 1	Ala 690
Ala	ьeu	mar	Arg	Phe	Sem	Rig	Leu	ΑĴa	A ì.a	Asp	Thr	IJe	Va1	Asn

PCT/US99/11743 WO 99/60986 Gly Ala Ala Thr Scr His Leu Ala Pro Thr Asp Pro Ala Asp Arg Leu Met Amp Thr Cys Arg Glu Cys Gly Ala Arg Ala Leu Glu Leu Val Gly Gln Leu Glu Asp Gln Thr Val Leu Arg Arg Ala Glu Pro Ser Lew Met Arg Ala Pro Lew Gln Gly Tle Lew Gln Lew Gly Gln Asp Leu Lys Pro Lys Ser Leu Asp Val Arg Gln Glu Glu Leu Gly Ala Met Val Asp Lys Glu Met Ala Ala Thr Ser Ala Ala Ile Glu Asp Ala Val Arg Arg Ile Glu Asp Met Met Ser Gln Ala Arg His Glu Ser Ser Cly Val Lys Leu Glu Val Asn Glu Ang The Leu Asn Ser Cys Thr Asp Leu Mct Lys Ala Ile Arg Leu Leu Val Met Thr Ser Thr Ser Lew Cln Lys Giu Jie Val Glu Ser Gly Arg Cly Ala Ala Thr Glm Glm Glm Phe Tyr Ala Lys Asn Ser Arg Trp Thr Glm B65 Gly Leu Tle Ser Ala Ser Lys Ala Val Gly Tro Cly Ala Thr Gln **B**5 Leu Val Glu Ser Ala Asp Lys Val Val Leu His Met Gly Lys Tyr **B95** Glu Glu Lou Tle Val Cys Ser His Glu I)e Ala Ala Ser Thr Ala Gin Leu Val Ala Ala Ser Lys Val Lys Ala han Lys Asn Ser Pro His Leu Ser Arg Leu Glm Glu Cys Ser Arg The Val Ash Glu Arg 

Ale Ala Asn Val Val Ale Ser Thr Lys Ser Gly Gln Glu Gln Ile

950

955

960

Glu Asp Arg Asp Thr Met Asp Phe Ser Gly Lou Ser Lou 11e Lys 965 970 975

Leu Lys Lys Gln Gln Met Gla Thr Gln Val Arg Val Leu 980 985 990

Glu Lyo Thr Leu Glu Ala Glu Arg Val Arg Leu Gly Glu Leu Arg 995 1100 1105

IMB Gln His Tyr Val Leo Ala Gly Gly Met Gly Thr Pro Scr Glu 1110 1115 1120

Glu Glu Pro Ser Arg Pro Ser Pro Ala Pro Arg Ser Gly Ala Thr 1125 1130 1135

Lys Lys Pro Pro Deu Ala Glo Lys Pro Ser Ile Ala Pro Arg Thr 1140 1145 1150

Asp Asn Glo Leu Asp Lys Lys Asp Gly Val Tym Pmo Ala Glo Leu 1155 1160 1165

Val Asn Tyr

- (2) INFORMATION FOR SEQ ID NO:12:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: other DNA
- (iii) JIYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 12:
- GAAGATACCC CACCAAAC 18
- (2) INFORMATION FOR SEQ ID NO:13:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other DNA
- (iii) HYPOTHETICAL: no

- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GCTTGACAGT GTAGTCATAA AGGTGGCTGC AGTCC 35

- (2) INFORMATION FOR SEQ ID NO:14:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: other DNA
- (iii) HYPOTHETICAL: no.
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 14:
- GGACATOTOC AGGGAGITGA ATAC 24
- (2) INFORMATION FOR SEQ ID NO:15:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 41
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: other nucleic acid
- (iii) HYPOTHETICAL: no.
- (iv) ANTI-SENSE: yes
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: humant
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CUACUACUAC UACUAGGECA EGEGTEGACT AGTACGGGII GGGIIGGGII G 41

- (2) INFORMATION FOR SEQ ID NO:16:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 516
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human

60 104



WO 99/60986			PC1	[/D <b>8</b> 99/11
(x) FEATURE: exon I of HIP)				
(xi)SEQUENCE DESCRIPTION: SEQ	ID NO: 16:			
TCTGTGGAAG GTTTGGAGGG GAGAGAGGGG		CECTIGGGCC	AUGGTOGOOD	60
CYGATCTCTC CGCCTCTTCC TCCTGCTCCC				120
CONTETETT GTGCGGGCTT TAATTGCCAT	GTTGTTGTGC	CAAGGGAGTG	AGTGGCGGCG	180
GGACCAGCAG CTGGCCACAC CCAATCCCAG				240
GCCACCTGGC TCCTGGGAGC GCTGCCCACC				300
CGACCACCCG TGGGCTGGG GCACCTTGGC				360 420
GGGTCTCAGC CACTCTCAGA GGCTTATTCA TTTTCAGACT GTCAGCATCA ATAAGGCCAT				420 480
ACACCCCAGA ANTATCCTTT GGATGTTGCT		dPD9100010	IMMOOMM	516
(2) INFORMATION FOR SEQ ID NO: (i) SEQUENCE CHARACTERISTICS:				
(A) LENGTH: 193				
(B) TYPB: nucleic acid				
(C) STRANDEDNESS: double				
(D) TOPOLOGY: linear				
(ii)MOLECULE TYPE: genomic DNA				
(iii) HYPOTHETICAL: po				
(iy) ANTI-SENSE: no				
(vi) ORIGINAL SOURCE:				
• •				
(A) ORGANISM: human				
(x) FEATURE: excm 2 of HIP1	m. No. 10			
(xi)SEQUENCE DESCRIPTION: SEQ		G)	1200000101	60
TOTTTTCCAT AACCCCCCT CACCGTGCAT GACCTTCTGC TCTGTTGTCA ACCGCCTGCC				
GROUPENER WINGSTEINE ACCOUNTED				180
CTATGGGGGGGGGA	200 10 0 1 E21221112 01		2110213000	193
01111110000				
(2) INFORMATION FOR SEQ ID NO:	:18:			
(i) SEQUENCE CHARACTERISTICS				
(A) LENGTH: 104	•			
• •				
(B) TYPE: nucleic acid				
(C) STRANDEDNESS: double				
(D) TOPOLOGY; linear				
(ii)MOLECULE TYPE: genomic DNA				
(iii) HYPOTHETICAL: no				
(iv) ANTI-SENSE: no				
(vi) ORIGINAL SOURCE:				
(A) ORGANISM: human				
(x) FEATURE: exon 3 of HIP1				
(xi)SEQUENCE DESCRIPTION: SEQ	ID NO: 1.8:			
The state of the s	-	********	(13 hmm) 11 11 11 11	6B

GTGTTCTTTT GCCCCTGCAG CTCCTGAAAG ACTCTCTGAG ATACAGAAAT GAATTGAGTG

(2) INFORMATION FOR SEQ ID NO:20:

ACATGAGCAG GATGTGGGTG AGTTTGGAGA TGTACTCAGG AGCC

(i) SEQUENCE CHARACTERISTICS:

(A)	LENGTH:	327

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FBATURE: exon 4 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AATTCCTGGC TCCACATCTC TTGACTGTTA TGTTCTTGTT CTTCACTCTC TTTCCCCTCC 60
TCTTCCTAAA AGGGCCACCT GAGCGAGGGG TATGGCCAGC TGTGCAGCAT CTACCTGAAA 120
CTGCTAAGAA CCAACATCCA GTACCACACC AAAGTGACTC TCTCCGCACA GTTCTCCCCC 180
CACCGGGGC TCCCCTGCTC CATCCCTTCA GCCCCTCCCT GGGCTCATTT GTCAGCTCTT 240
TCAGGTAATA GACACCCGAC CCTTCTCACC AAGTGCCCAC ATCATCTACC CAACCTCTCA 300
GAGAGGAAAG CCACCGCCAG GCCCACG

### (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 331
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) IXYPOTHETICAL: no
- (jv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exan 5 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGGCTCAACC ANTICITICICA CCTCGGCTC CCAAGTAGCT GGGACCACAG GCGTGTGCCA 90
CCACGCCCGG CTGAGAGAGG GCTCTTCATG TCTTCTGCCC TGACTCCCTT CCTCTCCCTC 120
CCTCCAGAA TCCCACCTTC CCAGGCAACC TGCAGATGAG TGACCGCCAG CTGGACGAGG 180
CTCGAGAAAG TGACGTGAAC AACTTGTAAG TGGCTCCTGC CCTGAGCCCA GGGAGGGAGA 240
AAGCTTTTGT GAATGCTGAC ACTTCTCATA ACCCTCATC ACCCCCTCAT GGGGGAGGAG 300
CGTGGCTGGG ATCCCCCACCA AAGCCCCTGGG G

### (2) INFORMATION FOR SEO ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 470
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no



# (vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(x) FBATURE: exon 6 of HIP!

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 22:

` '		•				
ACTOTCGCTG	TCACTGTTGA	CTTUACCAGG	CTCCATGGCC	ATAATACCCA	CAAGGCTAAG	60
ACTIGGACCI	GGAGTTCTCT	CTCTCTTTCC	CCATCCACAT	GACCATTGGA	GACTOGAGTA	120
CCCTAGAGCG	TGGGGGAGGG	GACAGGTAAC	AGACCGGCCT	CAGGCTGTGG	AGTGTAAGCT	180
CTCTTTCCTC	TTGGGTCCAG	TTTCCAGTTA	ACAGTIBUAGA	TGTTTGACTA	CCTGGAGTGT	240
CIAACITCAACC	<b>ፕሮ</b> ፐፕሮርልልልሮ	${\tt AGGTGAGTCT}$	CTTCCCTCCC	GTETAACCCA	GGCTCTCATG	300
GGAACTACCT	AATTOOTAGT	COTOCTOTO	CIGCAAAGTG	TGCAGGAGAA	GGGGTAGGAA	360
<b>ΑΑΤΟΙΓΑΙΚΑΣ</b>	TTCACACCCC	ATCTCTGGTC	TETCCAACCC	TOSTGONGOG	AGGGACTGAA	420
CCTCTTCAGT	APPTTTCTTT	TTAAGAGACA	AGGTCTCGGC	COCCTCCACT		470

### (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 565
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE; πο
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) PRATURE: exon 7 of HIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ć						
TCTTCACCTG	TTTAATGGGG	<b>ኢ</b> ዋልር:ሮተሞምእር	${\tt CTMTCTCMTG}$	GGAGTGTTGT	GAAGGTTAAL	60
ፕሮእስ፻亚አርስጥ	GAGGTAAAGC	ACGCACAGAA	TOSGTCCTTG	GIGTATGT2G	GACCCC10000	120
TCTGCCCCTC	TUANCACCC	$\mathcal{G}\mathcal{C}\mathcal{C}\mathcal{T}\mathcal{C}\mathcal{T}\mathcal{N}\mathcal{N}\mathcal{T}\mathcal{C}$	CCCTGGCTCT	$\lambda CC \lambda CC TT TC$	TCCCTCACTT	180
$\Delta \Delta V \Delta \Delta \Delta \Delta C C \Delta V$	GTATTCAACT	CCCTGGACAT	GTCCCGCTCT	GTGTCCGTGA	CCCCACCACC	240
GCAGTGCCGC	CTCCCCCCCC	TGATCCACCT	CATCTTGGAC	TYPEAGCCACC	TTTATGACTA	300
CACTOTORAG	CPTCTCTTCN	$\lambda \lambda CTCC\lambda CTC$	CTGTGAGTAC	CGCGGGCCAG	ATCTTCTTAC	360
ATGAGATTCA	BGCCAGAGGG	AGGATYCYCAG	CUTGAGGATG	TOCCCAGACA	AACCCACTCC	420
USIDACTOR	CTTTGGCTGT	CTCCTTCTCT	TOCANANGGO	CCCGGAGCTT	CTGACCATTG	480
<b>ፓርአር</b> ርአንንልንል	AGAGCAGGGC	CCAGGCTTTG	GTGACCCCAG	TARAGECCCT	CCCTTCCCAC	540
COTOCCTCC	ACTOTYACAG	GATCT				565

### (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 233
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FBATURE: exon 8 of HIP1

PCTYUS99/11743

WO 99/60986 PCT7	US99/111
(xi)SBOUENCE DESCRIPTION: SEQ ID NO: 24:	
GGGACAGGTG TARRECTATION OF CONTROLLED THE SECURITION OF THE TARRECT CAGGGTAGGT	60
GGGCCCTCC CCTCGAGAG CUCCGCTGTG GCTTCCCTGC CCTCTGGTCC CCCTCCCTC	120
TUACACTOTT TECANTITICT TEEACHECTC COMBETGACA DECUGGAAGG DUACCGGGAU	160 233
CGCTTCATGG AGCAGTTAC ARAGTRAGTG GTTCAAGTAA CAGGANTGGA GGT	<b>2</b> 55
(2) INFORMATION FOR SEQ ID NO:25:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 578	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(ii)MOLECULE TYPE: genomic DNA	
(iii) HYPOTHETICAL: no	
(iv) ANTI-SENSE: no	
(vi) ORIGINAL SOURCE:	
(A) ORGANISM: human	
(x) FEATURE: exons 9 and 10 of HIP1	
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 25:	
TGAATTCCAS CACCATGGAG TTTATCTCCT TGACACCCTC TCCCTTTCCC CTCCCCACCC	60
GGCACCANAG CENGGTGGCT SCTCTGTCCC CTACATGGGG CTGATGAAGA CACCCAGCAC	120 ายต
CCCTCAGGTO UTTUTUCACO CCTAGGTIGA AAGATOIGIT CTACCGCTDO ACCAACCTGC AGTACTICAA GCGGCTCATT CAGATCCCCC AGCTGCCTGA GGTAAGCATG CCCAACCACA	240
CACCTOGGO AUTGCAGAGE COCCAGGTAC TOTOTVAAGG GCCGCCCCC CCTCCCAAGC	300
AMAYAMTATT TGAGGATGTG TCTCCGTCTT CAGAACCCAC CCAACTTCCT GCGAGCCTCA	360
GCCTTSTCAG AACATATUAS CCCTGTGGTG GTGATCCCTG CAGAGGCCTC ATCCCCCCAC	420 4 <b>8</b> 0
AGCACTTOO AGAGAAACTT GGCCTTTCCT CTCACCTGCA AGTACAGGGG AGAGGCTGGG	540
GERGACCOTG GCCAAAGCCC ATTGACTCTA ACCAGGTT	578
(2) INFORMATION FOR SEQ ID NO:26:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 390	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(ii)MOLECULE TYPE: genomic DNA	
(iii) HYPOTHETICAL: no	
(iv) ANTI-SENSE: no	
(vi) ORIGINAL SOURCE:	
(A) ORGANISM: human	
(x) FBATURE: exon 11 of HIP1	
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 26.	60
ARRARATTI ARAAARITAA ACAUGTOTGA ACCUTTTAAT TOGAGAAAGO GGGCATTUTO UCATATGACT CAACTGACCO ACACACAGAA TICTOTGGCT CTCTGACTTA TICTOACTU	120
THTTTGGTCA ACCACAGAAT TTAITIGACA ACAAGITIGA TGACATCTTT GOCACTTCAT	טאנ
TOAGCAGTBA TCCCTPOART TYCAROAGTC AAAATGCTCT CAACAAGAAT GAGAAGTGAG	240
THE THE PROPERTY OF THE PARTY OF THE PROPERTY	300 360
CACTAACCAA AGAGGAATTC TTAATGATAC TGGGGCTTUT TAGATACAGA ACATCTTCAA	390
GGGTTGGGGG CAATGUUTTA IXCC/GTAAT	_ = -

300

360

420

436



### (2) INFORMATION FOR SEQ ID NO:27: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 547 (B) TYPE: nucleic soid (C) STRANDEDNESS; double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE; (A) ORGANISM: human (x) FEATURE: exon 12 of HIP1 (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 27: AAAATCAATA ACCATGGATT TATGAGTATT AGATTAGTAT CTGGTAACAT TTAGAGTATA -60 ATTINITIESCA TITCANACIAN ATGICCCONN ATTANTACION CUITTITANTI TOCCUCULUCTO 120 MCCTCACAAT TAAAAACAGA GGGATAGAAG CACTATGAAA GCAAACTCAT TCCCCTTCTC 180 240 TAGAAAACAT GAAGACTGAG GTATAACTTG GATCTGCTCT GCCTTTGCGC TTCACCAAAA 300 CACCOTAGAT TEGRATOTEN ANTITECNED ACACTACCON COCACACTOR CTCACACCOC 360 TARTECTAGE ACTITIGGAG GCCARGGCAG GAGGATTACE TGAGGTCGGG AGTTCGAGAC 420 CASCUTEGE AACAGGUGA AACCCCCCTC TTCAATAAA ATGCAATAAT TAGCCGGGTG 480 TSTTUGUAGO CACCTOPANT CCCAGCTACT CGGGAAGETE AGGCATGAGA ATTGCTTGAA 5**4**0 CTTGGGA 547 (2) INFORMATION FOR SEQ ID NO;28; (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 436 (B) TYPE: nucleic acid. (C) STRÁNDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 13 of HIP! (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 28: CCCCAGCCA CTCTARAGAG GACCACART CCCCGGCCCAT CATCCCCTGT TATTCTTGTT 60 ENTITAGES CTUUTAATGA COAGATOSTO CAACCOTOCT CCCACCTCCA GAOFTGACTT 120 AGGGGAATCA GGTATTTACT TGGANGACATG GTAGGACCG CTTCTCTGCCCCC CCATGCCCGT 180 GACCCGTGGC MGTGGGCGGT TGGCCTCATU ACCGGAGTCC CCCCACAGAG CCACXCGGTT 240

GTGCTGCAGC TEAAGGGCA GGCGAGAGCGCC ACCCCCAC ACCCCCCA GCACCAGCAC

CICCGCCAC ACCCCCCC CONCIPERA TICCIPERES CAGAACTOCA CENCTUAGG

AGGCAGCGGG AGGACACCGA GAAGGCTUAG CGGAGCCTGT CTGACATACA AACACACCCC

- (2) INFORMATION FOR SEQ JD NO:29:
- (i) SEQUENCE CHARACTERISTICS:

TGGGTGGGGG CGGGGG

- (A) LENGTH: 469
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 14 of HIP1

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 29:

-				
AAGCCTTCCAG	TGAACACTGA	ͲͲϢͲϾϾϾϔ	CCACHILLIAGO	60
GTCTCAAAAC	AAAACAAGGA	GGACCTICTA	GGGACCCTGG	120
ACTICIO:8GCT	ACCTTAGACT	CCTCACCTTC	CTCCTTTACA	180
DDTCAACACC	CATATACTAR	GCTAAAGGAG	AAGTACAGCG	240
3231 GROUP (AUDIO	OCCAROVEDING:	ስተተቀቀጥሮ <b>አ</b> ርተርተ	CONCINCACON	300
GACCIGCIGC	DISAMOGIANO	9000103300	ancececer:	360
TAGGGAGACA	GCGGCTCMGG	CCLANCASCAL	Ct.L.C.GGGGCC	
CTAAGGCATT	GCCGTCATUT	CGGGAAUCAC	acutttuas	425
				459
	GTCTCAAAAC GGTCCCGGCT AATGAACAGC GACCTGCTGC TAGGGAGACA CTAAGGCATT	GTCTCAAAAC AAAACAAGGA GGTCCCTGCT BUUTTACACT AATGAACAGC GATATAGCAA GACCTGCTGC BUAAAGGTAAG TAGGGAGACA GCGGCTCAGG CTAAGGCATT GCCGTCATCT	GTCTCAAAAC AAAACAAGGA GGACCTTCTA GGTCCCTGCT AGCTTAGACT CCTCACCTTC AATGAACAGC GATATAGCAA GCTAAAGGAG GACCTGCTGC GGAAGGTAAG ACCCTCAGCC TAGGGAGACA GCGGCTCAGG CCTGTGGCTT	AAGGCTGCAG TGAACACTGA TTCTCCCACT CCACCCCAGC GTCTCAAAAC AAACAAGGA GGACCTTCTA GGGACCCTGG GGTCCCTGCT BEUTTAGACT CCTCACCTTC CTCCTTTACA AATGAACAGC GATATAGCAA GCTAAAGGAG AAGTACAGCG GACCTGCTGC BGAAGGTAAG ACCCTCAGCC CCTGTCACCA TAGGGAGACA GCCGCTCAGG CCTGTGGCTT CCCCGGGGCC CTAAGGCATT GCCGTCATCT CGGGAAUCAC ACUTTTCAG TEGGCTGTGT CCTGCGTCCC AACCCATC

# (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 359
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS; double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: buman
- (x) FEATURE: exon 15 of HIPL

WASHOURNOE DESCRIPTION: SEO ID NO: 30:

(メリンなん) いいい	ふかい ひじりかいかい	LINOIS PEÓ	10. Jul			
CCGTAGGAAA	GTGATTCCTG	TGTCTGACTC	TAGGGCACCC	ACACCCTGAG	TATGATTGTC	60
CW0 IS 0 91313 913	CARCERNATION	AACCCTCCCA	TOTOCTEGTT	CANGACACTG	TTCTTCTT!'	120
	CACCTCACCA	AACACCTCTC	CATGGGGGAGA	CARGOCCAGG	TAGATTTCCA	180
GUNGANTUKA	CAUCHONILIS	101CADDIGIC	447400000000T	እረጥሮአሮሮአፍር	GCCAGCGGAA	240
acgagaaaa	AAAGAGCTGG	AGGATTCGTT	GGMGCGCMIC	acamaman news	mest were the following	300
GGTGAUTGGG	ACGAGGAGCA	CACCCCCAVVA	GMARIAAGGGG	GCTGTTGAGT	TGGTGGGGG	359
CCCTTTGTGG	CCTTCTGCTC	CATGGGCAGT	TCTGTGGGTC	GGMGGGGATC	AGAHABUAG	723

## (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 209
- (B) TYPE: nucleic soid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no

4 B 5

WO 99/60986	
A RESULT CUETO SECUTO	

WO 89MH9R6	PCT/US99/1174
(iν) ANTI-SENSE; πο	
(vi) ORIGINAL, SOURCE:	
(A) ORGANISM: human	
(x) FEATURE: exon 16 of HIPI	
(xi)SEQUENCE DESCRIPTION; SEQ ID NO; 31;	
GTTGATCCCT TCGGACGTTT TTACATTTTT ATATTCTTTG TCACTGTCAC CCAGATCA	GA 60
STECCTOTOT TTTTCTTUTC TIPOAGACTO AAGAACAGUT GGAAGTTUTA GAGAGCTT	
AGCAGGAACT TOCCACAGG CAACGGGAGC TTCAGGTTCT GCAAGGCACK: CTGGAAAC	TT 180
UTCCCCAGGT AAATACCTCC TITTTTTTTT	209
(2) INFORMATION FOR SEQ ID NO:32;	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 485	
(B) TYPB; nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
• •	
(ii)MOLECULE TYPE: genomic DNA	
(iii) HYPOTHETICAL: no	
(iv) ANTI-SENSE: no	
(vi) ORIGINAL SOURCE:	
(A) ORGANISM: human	
(x) FEATURE: exon 17 of HIP1	
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 32:	
CCCCCACTGC AATCAGTGTG TCCCCGGGAG GGAATCAGAG TGGCAGGTTA AAGAGCCA	rc 60
ACCTIVICIAG TECTTECARE CEGGICGICA CPTENACETE INXIGAAGIAG GEACTGTT	
ACTCAACCAG COTCTCCCTC TTTCCTTGTG GTCACCTTTG CAGTCAGAAG CAGACTCCC	
AGCINANTEC GCCGAGCTAG AGAAGGACIN: CHACACCCTC CTCAGTGGCG CAGCTCATI	
GGAGGAGGAA TTATCTCTC TTCGGAAAGA AUTGCAGGAC AUTCAGCTCA AACTYGCCC	
CACACACIGET CACGGACATE GACACGAGOS ACCACCTOTO AATTCCCACC GAGGGCCTY	
CERCATGUAG GGAGGUTGGG ACCACCEGG GGUTGUTGAG AAGGGUTTTG GGGCCTTXX CTGATTGTGC ACACATTGTG TAGGTGTAAT GCCAGCAGAG CUTGCATTGC CTCCACAC	
TIGHTIGIGE NOW, ATTEME TAMESTSTANT GUING AGGO CUTGUNTTEC CHUNCAC	N: 480

- (2) INFORMATION FOR SEQ ID NO:33:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 468

CATCA

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) PEATURE: exon 18 of JIIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 33:
- TTACTOSCTT GEACCTCATT GOOCATGACT TGAGCTAAGA TGCTAAGAGC CCCAGCCAGG 60 TCATCCTGCT CAGGITCATT ATGGAGTCTA GGGCAGACTC TCACCTCCCT GGACCATTTT 1.20

WO 99:60986  ***ROSSETUCIOSA GEGARGATE GEGARAGACE NACONANAT GETTETGGGG GEGETECTORA GEGETECTORA CARGEGECET TORANAGECET TURANAGECET GEGARGAGE CARGEGECET GEGARGAGE CARGEGECET GEGARGAGE CARGEGECET GEGARGAGE CARGEGECET GEGARGAGE CARGEGECET GEGARGAGE CARGEGECET GEGARGAGE CARGEGECATE CARGETTARCE TRANSCATURA SETTEMBOSE DITTERARCE 300 ARCHITECTUR CARGETTARCE TRANSCATURA SETTEMBOSE DITTERARCE 300 ARCHITECTUR CARGETTARCE TRANSCATUR SETTEMBOSE DITTERARCE 300 ARCHITECTUR CARGETTARCE CARGEGECE CARGEGECEA TECTAGRACA GITACATER 301 ACCURATION FOR SEQ ID NO:34:  (3) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 393  (B) TYPE: nucleic acid  (C) TOPOLOGY: linear (i)MOLECULE TYPE: genomic DNA (ii) HYPOTHETICAL no (ii) ANTI-SENSE: no (ii) OF HIPI (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 34: CAPTAGTAR CONCRETA CARGEGETTAR TRANSCAGE CARGEARAGE ACCURAGAGE GETAGRACA CONCRANAGE CARGEARAGE ACCURAGAGE TRANSCAGE CARGEARAGE ACCURAGAGE ACCURACAGE ACCURAGAGE ACCURAGAGE ACCURACAGE ACCURACAGE ACCURACAGE ACCURACAGE ACCURACAGA ACCURAGAGE ACCURACAGA ACCURACAGA ACCURA			
ASSETTICION GENERALIZA CANAGERET TRABASACCE CONTENTA 240 GENERALIZA CANTONICA TRABASCEC CANAGERET TRABASACCE CONTENTA 300 ASCETTARRA CANTAGEM TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTAC. TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTAC. TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTAC. TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTACACT TRASASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTACACT TRASASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTACACT STRATITICAT ANTITUCTT  (2) INTORMATION FOR SEQ ID NO:34: (3) SEQUENCE CHARACTERISTICS: (4) LENGTH; 393 (B) TYPE; nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY; linear (ii) MOLECULE TYPE; genomic DNA (iii) HYPOTHETICAL to (iv) ORIGINAL SOURCE: (A) ORGANISM; human (A) FEATURE: exon 19 of HIP1 (A)SPQUENCE DESCRIPTION: SIQ ID NO: 34: CONTRASASCE CONTRACAGEA CONTRACAGE TRACAGEAC TOTOCORGAN CONTRACAGEA CONTRACAGEA TRACAGEACAC TOTOCORGAN CONTRACAGA CANAGEACAC TRACAGEAC TOTOCORGAN CONTRACAGA CANAGEACACT CONTRACAGA CONTRACAGE CONTRACAGA CONTRACAGA ANGSENTOR OCCARACTEM TOTOCORGAN CONTRACAGA CONTRACAGA ANGSENTOR OCCARACTEM TOTOCORGAN CONTRACAGA CONTRACAGA ANGSENTOR OCCARACTEM TOTOCORGAN CONTRACAGA CONTRACAGA CONTRACAGA CONTRACAGA CARGAGAC CONTRACAGA CONTRACAGA CONTRACAGA CONTRACAGA CARGAGAC CONTRACAGA CONTRACAGA CONTRACAGA CONTRACAGA (C) STRANDEDNESS: double (D) TOPOLOGY: linear (II) MOLECULE TYPE: genomic DNA (III) HYPOTHETICAL TO (IV) ORIGINAL SOURCE: (A) ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 o	WO 99/60986	PC	77/0899/11743
ASSETTICION GENERALIZA CANAGERET TRABASACCE CONTENTA 240 GENERALIZA CANTONICA TRABASCEC CANAGERET TRABASACCE CONTENTA 300 ASCETTARRA CANTAGEM TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTAC. TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTAC. TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTAC. TRABASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTACACT TRASASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTACACT TRASASCEC CONSCICLA POCTARRAC STRACAGEC ASCARCATTY CARCATTACACT STRATITICAT ANTITUCTT  (2) INTORMATION FOR SEQ ID NO:34: (3) SEQUENCE CHARACTERISTICS: (4) LENGTH; 393 (B) TYPE; nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY; linear (ii) MOLECULE TYPE; genomic DNA (iii) HYPOTHETICAL to (iv) ORIGINAL SOURCE: (A) ORGANISM; human (A) FEATURE: exon 19 of HIP1 (A)SPQUENCE DESCRIPTION: SIQ ID NO: 34: CONTRASASCE CONTRACAGEA CONTRACAGE TRACAGEAC TOTOCORGAN CONTRACAGEA CONTRACAGEA TRACAGEACAC TOTOCORGAN CONTRACAGA CANAGEACAC TRACAGEAC TOTOCORGAN CONTRACAGA CANAGEACACT CONTRACAGA CONTRACAGE CONTRACAGA CONTRACAGA ANGSENTOR OCCARACTEM TOTOCORGAN CONTRACAGA CONTRACAGA ANGSENTOR OCCARACTEM TOTOCORGAN CONTRACAGA CONTRACAGA ANGSENTOR OCCARACTEM TOTOCORGAN CONTRACAGA CONTRACAGA CONTRACAGA CONTRACAGA CARGAGAC CONTRACAGA CONTRACAGA CONTRACAGA CONTRACAGA CARGAGAC CONTRACAGA CONTRACAGA CONTRACAGA CONTRACAGA (C) STRANDEDNESS: double (D) TOPOLOGY: linear (II) MOLECULE TYPE: genomic DNA (III) HYPOTHETICAL TO (IV) ORIGINAL SOURCE: (A) ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 of HIP1 (XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (ORGANISM: human (X) FEATURE: exon 20 o	TAGAATCTAT GYGCCAGCTY GCCAAACACC AACGAAAAAT GCTTCTGGTG	GGGTCCNGGN	
ACCITAGAC TETETANT TEARTET A TOTAGACT A SCIENCISC STOCKAGES STOCKAGES ACCIDENT CAGAMITAC TRADEGICC CLASSICCA TECTAGACA TRACATETS 420 468 468 468 468 468 468 468 468 468 468	ACCOMPRIES GEAGGEGATA CARGAUGUCU TGAACCAGUT TGAAGAACCT	CCTCTCATCA	<b>-</b>
### ACCOUNT OF THE PROPERTY OF	GETGEGETGG GTCTGCACCT ACACTTGCAA TTCCCCACCT GGCAGGGGCC	AGGTCCTTAC	
(2) INFORMATION FOR SEQ ID NO:34: (3) SEQUENCE CHARACTERISTICS: (A) LENGTH: 393 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (i)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIP1 (xi)SEQUENCE DESCRIPTION: STQ ID NO: 34: cachaghage characterist characterist thangage transpared characterist characteris chara	ACCOMMAGAC TOTATTGATG TREATOTCA TOTGAGACTT ACCTUAGGOS	CTC3CAGCCC	
(2) INTORMATION FOR SEQ ID NO:34; (3) SEQUENCE CHARACTERISTICS; (4) LENGTH: 393 (B) TYPE; nucleic acid (C) STRANDEDNESS; double (D) TOPOLOGY; linear (ii)MOLECULE TYPE; genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (v) ORIGINAL SOURCE; (A) ORGANISM; human (x) FEATURE; exon i9 of HIP1 (xi)SPQUENCE DESCRIPTION; SEQ ID NO: 34; cactragrams creetgeart characterian transpaces transparant exceptions transparant exceptions are exceptional experimental expe	AGCAGCATCT CAGCATTACC TIMESESCEC CONSCIUNCA TECTMONIES	GTTACATOTO	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 393 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIP1 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34: cactracteriage chocytegrath chargesteria thangeneral trunscence truckering generation of the truckering generation of the truckering generation of the truckering generation of the truckering generation of truckering ge	CAMACACAGA GCALLAGAGC CININCKCIN GINIIII1901 NIII1011		
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 393 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIP1 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34: cactracteriage chocytegrath chargesteria thangeneral trunscence truckering generation of the truckering generation of the truckering generation of the truckering generation of the truckering generation of truckering ge			
(A) LENGTH: 393 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIP1 (xi)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarg criccifcart cartectita thangarga tuangocage transparet (si)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarg criccifcart cartectita thangarga tuangocage transparet (si)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarg criccifcart cartectita thangarga transparet (si)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarge criccifcart and criccifcart and security and criccifcard transparet (si)Agratage criccifcart acternaria criccifcard assectation and security and secu			
(A) LENGTH: 393 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIP1 (xi)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarg criccifcart cartectita thangarga tuangocage transparet (si)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarg criccifcart cartectita thangarga tuangocage transparet (si)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarg criccifcart cartectita thangarga transparet (si)SPQUENCE DESCRIPTION: SIRQ ID NO: 34: cartagtarge criccifcart and criccifcart and security and criccifcard transparet (si)Agratage criccifcart acternaria criccifcard assectation and security and secu	(i) SBOUENCE CHARACTERISTICS:		
(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (X) FEATURE: exon 19 of HIP1 (Ai)SFQUENCE DESCRIPTION: SBQ ID NO: 34: CACTAGTAGE CICCITECATO CAGTECTTAR TTAGGAGGA TUAAGCCAGC TATCMARACT TOPCEGET CATCCACCAA CTEGRAGAAA GCTGGAGGA TUAAGCCAGC TATCMARACT TOPCEGET CATCCACCAA CTEGRAGAAA GCTGGAGGA TEGRAGGAG TECAGAGAG GFARGAATGG CAAGGAGAA CTEGRAGAAA GCTGGAGGAA GCTGATACCAT GCTCACCAGAG GFARGAATGG CAAGGACAA CTEGRAGAAA GCTGCACAGAA GCTCACACACAA CCTGATAGGG GGAATGAGAA ACTGCTCACCA TCC ACAGCAGACC TGAGAAACTT CTCTTTCCAA TCC 300 360 373  (2) INFORMATION FOR SEQ ID NO:35: (i) SEQUENCE CHARACTERISTICS: (A) LENCIT: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (iv) ORIGINAL SOURCE: (A) ORGANISM: buman (x) FEATURE: con 20 of HIP1 (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 35: GCCTGCCCA GAAGGTAGAA ARGCCAAGGA CACACTCTA TUGGCTAGAC AAGGACCTAA TGTAGGCAAC TCATACAGA ARGCCAAGA ACTCTCTT TUGGCTAGAC GAGAACCTA TGTAGGCAAC TCATACAGA ARGCCAAGA ACCACTCTA TUGGCTAGAC GAGAACCTA TGTAGGCAAC TCATACAGA AGCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACTTCAC AGCCTGAAA ACCACTCTCT TUGGCTAGAC GAGAACCTA TGTAGGCAAC TCATACAAGA AGCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACC TCATACAACA GACCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACC TCATACAACA AGCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACC TCATACAACA AGCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACC TCATACAACA AGCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACCC TCATCCAACAC AGCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACCC TCATCCAACAC AGCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACCC TCATCCAACAA AGCCCTGAAA ACTCTCTTT CCAATCCTGC TGGAGACCC TCATCCAACAA AGCCCTCAAACCA GACCTGACAA TCTGCAACTCTACATCTACAGA AGCCCTCAAACCA GACCTCTCCTACA TTGGCCAATCTACAGA TGGAGACCTTAAA CCCTCCCAACACCA CCCTCCTCCACAC TTGGCCAATCTACAGA TGGACTCCAACACCAACAA CACCTCCAACACCAACAA TTGGCCAACATCTACAAC TGGACTCCAACATCTACAAC AGCCTCCAACAACCAACAACAACAACAACAAACAAACAAA	*,		
(C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIP1 (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 34: CACTAGRARG CTCCTCCART CAGRECTERA TRANGAGGA TUAAGCCACC TATGREAACT CACTAGRARG CTCCTCCART CAGRECTERA CACARACCT CCCCTCCACG GTCACATCCA TOTCCAGCTC CANCORCAA CNGRAGAAA GCTCACATCCA CAGACAGGG TTCAGAAGCA CCTGRARGGG GGSATAGTGA CAGGCCCCC TECATCCA GTCACATCCA CCTGRARGGG GGSATAGTGA CAGGCCCCCT TECATCAGGA ACGCACACCAT CAGACAGACC TGCAGAACT CTCTTTCCAA TCC  (2) INFORMATION FOR SEQ ID NO:35: (3) SEQUENCE CHARACCTERISTICS: (4) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (vi) ORIGINAL SOURCE: (A) ORGANISM: button (x) FEATURE: exon 20 of HIP1 (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35: OSCCTGCCCA GAAGGGTAATA ACCCCAAGG ACACCCTCT TUGCCTACTC TUGCCTACTC AAGAAAMSCA TCTACGCAAC TCATCACAA ACCCCAAGG ACACCCTCT TUGCCTACTC CAGACACCTC AAGAAAMSCA CCCCTGCCCA GAAGGGTAAGA ACCCCCAAGA CACACCCTCT TUGCCAACCCACC CACACCACCC CACACCACCACC ACGCCTCCCA GAAGGGTAAGA ACCCCCAAGA ACCCCCACACACCACC CACACCACCACC CACACCAC			
(II) MOLECULE TYPE: genomic DNA (III) HYPOTHETICAL: no (IV) ANTI-SENSE: no (IV) ANTI-SENSE: no (IV) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIPI (xi)SPQUENCE DESCRIPTION: SBQ ID NO: 34: CACTAGTAGE CICCTCCATT CAGGECTTAR TRANCAGGE THAGGECAGE THATCAGAGCAGE TOTTCCAGCT CAGCCACCAA CAGGAGAAAA GETAGAGCAGE CAGGACAGGA TRACCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			
(ii)MOLECULE TYPE; genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (iv) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIPI (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 34: CACTAGTAGA CICCICCATA ACCESCRATE CACAGACACT CICCICCACG GTCACATCCA 126 TOTICCACCT CACCCAGGAA CIGGAGAAAA GCTGAGGCUA CICCICCACG GTCACATCCA 127 TOTICCACCT CATCCAGGAA CIGGAGAAAA GCTGAGGCUA CICCICCACG GTCACATCCA 128 CIDAGAATGG CCAGGACAGA CIGGAGAAAA GCTGAGGCUA CICCICCACGG GTCACATCCA 129 TOTICCACCT CATCCAGGAA CIGGAGAAAA GCTGAGGCUA CICCICCACGG GTCACATCCA 120 CCTGARTGCG GGGATAGTAA CACCTCCCTC TECACCAGAAAAGGCA CCCAACTCAT 120 CCTGARTGCG GGGATAGTAA CACCTCCCTC TECACCAGAAA 120 CCTGARTGCG GGGATAGTAA CACCTCCCTC TECACCAGAAA 120 CCAGACAGGC TGAGAAACTT CICTTCCAA TCC 130 CCAGACAGGC TGAGAAACTT CICTTCCAA TCC 130 C2) INFORMATION FOR SEQ ID NO:35: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (iv) ORIGINAL SOURCE: (A) ORGANISM: buman (x) FEATURE: exon 20 of HIPI (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35: CCCCTCCCCA GAAGGTAGGA ACCCCCAGG ACACCTCTG CCCCCCCCCA AAGAAAGSCA 120 ACGCTTCACA AACACCTCAA TACCCCAAGG ACACCTCTG CCCCCCCCCACAACACCACACC	· · ·		
(iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 19 of HIP1 (xi)SPQUENCE DESCRIPTION: SBQ ID NO: 34: CACTAGTARG CICCICCATT CAGTECTTA TTARCAGGA TUAAGCCAGG TATCAGAACT TOTCCAGCT CATCAGGAA ACCECCTTCA CACTAGTACT COTCCAGAGG TOTCCAGAAGAACT CTTCCAGCTG CACCAGGAAAA GETGGGGCAA GETGACATCCA TOTCCAGAGG CAAGGAAAA GETGGGGCAA GETGACATGCA CACAACTAAT ACAAGAAAGG CACAGGAAAACTT CICTITCCAA TCC (2) INFORMATION FOR SEQ ID NO: 35: (3) SEQUENCE CHARACTERISTICS: (4) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (X) FEATURE: exon 20 of HIP1 (Xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35: OGCCTGCAG AGGGAAAGGA CACACTCAT TAGACAACCAAC CACACAGACA CACACACAA CACACACAACAACCAACCA			
(iv) ANTI-SENSE: no  (vi) ORIGINAL SOURCE:  (A) ORGANISM: human  (x) FEATURE: exon 19 of HIP1  (xi)SPQUENCE DESCRIPTION: SEQ ID NO: 34: CACTAGGRAG CICCICCATT CAGGGCTTA TTAGGAGGA TUAAGCCACC TATCAGAACT  TOTCCACCT CATCCACCA CTGGGGAAA GCTGGAGCA CTGTCCACCAC TATCAGAACT  TOTCCACCT CATCCACCA CTGGGGAAAA GCTGGAGCAC CTCTCCACCAC TATCAGAACA  COTGATGGG CAAGGACA CTGGGGAAAA GCTGGAGCAC CTCTCCACCAC TATCAGAACA  CCTGATGGG GGGATAGICA CTGGGGAAAA GCTGGAGCAC CAGACAGGG TTCAGAACCA  CCTGATGGG GCAATAGICA CACCACCAT TACCACACACACACACACACACACACA			
(vi) ORIGINAL SOURCE:  (A) ORGANISM: human  (x) FEATURE: exon 19 of HIP1  (xi)SEQUENCE DESCRIPTION: SIRQ ID NO: 34: CACTAGTARG CTCCTCCATT CACTGCTTAR THACGASCA THACGCAGG TAMCACATCAT TOTCCAGCAG CTCCTCCATT CACTGCTTAR THACGASCA THACGCAGG TAMCACATCAT TOTCCAGCTE CAMCACCAA CTCCTCCA CACAMACACT CCTCTCCAGG GTCACATCAA TOTCCAGCTG CAMCACCAA CTCCTCAGG CACAMACACT CCTCTCCAGG GTCACATCAA TOTCCAGCTG CAMCACCAA CTCCTCTCAGG CACAMACACT CTCCCCAGAAG GTARACAAGG CAAGGACAG TCTCTCTCCGG CTACTMATAG CACACAGGG TTCCAGAAGAG ACCAGAAGG CATCTAGGCA ACCACTAAAA CGGGAGGACA CCCTATGAAA GTGTCACCAT ACACCAGACC TEAGAAACTT CTCTTTCCAA TCC  (2) INFORMATION FOR SEQ ID NO:35: (3) SEQUENCE CHARACTERISTICS: (4) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (i)MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 20 of HIP1 (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35: CGCCTGCCA GAAGGTAAGA ACCCCCAGA TECCGAGGAC CCTCTGCATC AAGAAAGGCA TCCTAGGGAAC TCATACAAGA AAGGCATCTA CCAACTCAT AAAACGGAG GAAGAGGTAT CAAGGGAAC TCATACAAGA AAGGCATCTA CCAACTCAT AAAACGGAG GAAGAGGCTAT CAAGAGGTCAA CCACCTCAA TECGGGGAA ACCTCCTTC CAACTCATC AAGAAAGGCA TGCTACACC TCCTCAGACCA GACCTCCAAA ACCTCCTCTC CCACTCCACC CACACCACA TGCTACACC TCCTCTAGAC AAGCCACCAA ACCTCCTCTCTC TCCACTCCAC			
(A) ORGANISM: human  (x) FEATURE: exon 19 of HIP1  (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 34: CACTAGTARG CTCCTCCATT CAGTGCTTAR TTANCGAGGA THACGAGACT TATCAGAACT GCCTCTCACC TTGCCCTGTT CAGTGCTTAR TTANCGAGGA THACGAGACT TATCAGAACT GCCTCTCACC TTGCCCTGTT CAGTGCTTAR TTANCGAGGA THACGAGACT TATCAGAACT GCCTCTCACC GTCCACCCACCACCACCACCCACCCCCCCCCC			
(x) FEATURE: exon 19 of HIP1  (xi)SPQUENCE DESCRIPTION: SBQ ID NO: 34:  CACTAGTARG CTCCTGGATT CAGGGCTTAR TRACGAGGA TGAAGCCAGG TATCTAGAACT  TGCTCTACC TTGCCCTGT PIVOCTCTCA CAGATAACT CCTCTCCAGG GTCAGATCA  TTCCCAGCTG CATCCAGGAA CTGGAGAAAA GCTGGAGGAA GTACCTGCCAGAAG  GTAAGAATGG CCAAGGACAG TCTUTUTGG CTACTCAGAA ACGCAACGGAT  CCTGAATGG GGGATAGTCA CAGGTCCCTC TGCATCAGA AAGGCAACTCAT  ACAGCAGACC TGAGAAACTT CTCTTTCCAA TCC  (2) INFORMATION FOR SEQ ID NO:35: (3) SEQUENCE CHARACTERISTICS: (A) LENGTH: 421  (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: genomic DNA  (iii) HYPOTHETICAL no  (iv) ORIGINAL SOURCE: (A) ORGANISM: burden  (x) FEATURE: exon 20 of HIP1  (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GGCCTGCCCA GAAGGTAAGA ACGCCCAAG ACACTCCTC TCUGCCTACTC AAGAAAGGCA  AGGGTTCAGA AGCACCTCAA TGCGGGAGTA GTGACAGGCC CCTCTCCATC AAGAAAGGCA  TGTAGGGAAC TCATACAAGA AAGCCCTACAA ACTTCTCTT CCAATCCATC AAGAAAGGCA  TGTAGGGCAAC TCATACAAGA AAGCCCTACAA ACTTCTCTT CCAATCCTCC CACACTCATC  TGGACTCCC CATCCAACAA GCCCTCCATAA CCCTCGCTGG CCACTTCATC CAATCCTCC CACACTCAGC  TGCACACCC CATCAACAA GCCCTCCATAA CCCTCGCTGG CCACTTCATCAC AACACAGCCA TGCCTATGG  TGCACACCC CATCAACAA GCCCTCCATAA CCCTCCTCGCC CACACTCATC CAATCCATCAG  TGCACACCCC CATCAACAA GCCCTCCATAA CCCTCCCCGAC TGCTCTCATCAC CACCACTCAGC  TGCACCACCC CATCAACAA GCCCTCCATAA CCCTCCCCGAC TGCTCTCTCTC CATCCCATCAG CCCCTCCCCAAC TGCCTCATCAG CCACTCCATCAG CCCCTCCCCAAC TGCCTCATCAG CCCCTCCCCAAC TGCCTCATCAG CCCCTCCCCAAC TGCCCAATCAG CCCCTCCCCAAC TGCCTCCCCAC TGCCTCTCCC CACCACTCAG CCCCTCCCCAC TGCCTCACCAC CACCTCAACC ACCCTCCCAC CCCCTCCCCAC TGCCTCTCCC CACCACTCAG CCCCTCCCCAC TGCCTCCCCAC TGCCCAC TGCCCAC TGCCCAC TGCCCAC TGCCCAC TGCC			
(xi)SFQUENCE DESCRIPTION: SIRQ ID NO: 34:  CACTAGTARG CTCCTCCATT CAGGECTTAR TRACCASCA TRACCACCACC TATCACAACT  GCCTCTGACC TIGCCCTGTG TRUCCTCTCA CACATACACCT CCTCTCCACG GTCACCAACCA  TTTCCAGCTG CATCCACCAA CREGAGAAA GCTGGAGUAG GTCACCAGAAG  TTTCCAGCTG CATCCACCAA CREGAGAAA GCTGGAGUAG GTCACCAGAAG  CCAGAGAAGG CACCAGGACA TTCTUTUTGG CTACCAGAAGG TTCAGAAGGA  CCAGAGAAGG CACCACCACA ACCCATAGAA ACGCACCACTA  ACACCAGACC TCAGACAA ACCCATAGAA CCCGACCACA ACCCACCACA  (2) INFORMATION FOR SEQ ID NO:35:  (3) SEQUENCE CHARACTERISTICS:  (4) LENGTH: 421  (B) TYPE: mucleic acid  (C) STRANDEDNESS: double  (D) TOPOLOGY: linear  (ii)MOLECULE TYPE: genomic DNA  (iii) HYPOTHETICAL: no  (iv) ORIGINAL SOURCE:  (A) ORGANISM: burian  (x) FEATURE: exon 20 of HiP1  (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GGCCTGCCCA GAAGGTAAGA ACGCCCAGG ACACCCACCA CCCCTCCACACA AGAAAGGCA  TGCAGGCAAC TCATACAAGA AGGCCATAA GTGACACCACAC CCCTCTCCATC AAGAAAGGCA  TGCAGGCAAC TCATACAAGA AGGCCACAA ACTCTCTT CCAATCCTTC CAACACCACAC ACCACCCAC			
CACTAGTARS CTCCTCCATT CAGGCTTAR TRACCAGER TRACCACC TATCCACACC TGCCCTCTCC TGCCCTGTG THEORETTER CACACACACA CACCACACA CTCCTCCCAC TATCCACACA CTCCCAGACA CCCCTCAGACA CCCACCAGACA CCCTCAGACA CCCACCAGACA CCCTCAGACA CCCACCAGACA CCCTCAGACA CCCACCAGACA CCCTCAGACA CCCACCAGACA CCCTCAGACA ACCCACCACACA CCCCTCAGACA CCCCTCAGACA ACCCACCACACA CCCCTCAGACA CCCCCCAGACACA CCCCTCAGACA CCCCCCAGACACA CCCCTCAGACA CCCCCAGACACA CCCCTCAGACA CCCCCCAGACACA CCCCTCAGACA CCCCCCAGACACA CCCCCAGACACA CCCCCAGACACA CCCCCAGACACA CCCCCAGACACA CCCCCAGACACACAC			
TECTOTERIC TECCOTOR TROCTOTA CACATACT COTTENERS 120 THECRACITE CATEGRICAN CHEGRICANA CUTGRAGULA GUTGRAGULA GUTGRAGULA GUTGRAGULA GUTGRAGULA GUTGRAGULA GUTGRAGULA GUTRAGARGA 240 CARARATGO COMBORA CHOUTUNG CHARACTERIS TOCATOMATA COCCANCTOM ROMAGRAAGG CATCITAGGA ACTOMARAA COGGARGAGA COMTATGANA GUTCACCAT  CAACCAGACC TEAGAACT UTCTTTCCAN TCC  (2) INFORMATION FOR SEQ ID NO:35: (3) SEQUENCE CHARACTERISTICS: (4) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 20 of HIPI (xi) SRQUENCE DESCRIPTION: SEQ ID NO: 35: GECTECUCA GAAGGTARGA ANGECCAAGG ACACTECTE TUGGCTACTC ATCGCCAGAC AGGGTTCAGA ACACCTONA TREGGGGATA GURACAGGTC COUTCTGCATC AAGAAAGGUA CARACGAGAC TOATACAAGA AAGGUATGTA COCANCTOTA GAAGGTGTCAC CATCACTCAGG TEGRACTIC CATTCCATAA CCCTGCTGGC CACCTTGACC TIGGCTAGAC TEGRACTIC CACCTAGACC AGCCTAGACA ACTCTTCTT CACATTCCTT CACATCCTTC CACCTACTAGAC TEGRACTIC CACCTAGACC AGCCTAGACA ACTCTTCTT CACATTCCTT CACATCCTTC CACCTACTAGAC TEGRACTIC CACCTAGACC AGCCTAGACA ACCTAGACCA TIGGCCTAGAC TEGRACTIC CACCTAGACCA GACCTAGACA ACCTAGACCA TIGGCCTAGACA CACCAGGCCA TIGGCCTAGAC TEGRACTIC CACCTAGACA CACCTAGACA CACCAGGCCA TIGGCCTAGACA CACCAGACCA TIGGCCTAGACA CACCAGGCCA TIGGCCTAGACA CACCAGGCCA TIGGCCTAGACA	(XI)SHQUENCE DESCRIPTION, DIAZ 112 1877. FF.	: тапсаєваст	60
TITICAGUTE CANCAGGAA CITGERGAAAA GUTGERGAAAA GUTARCAGGG TITCAGAAGGA 240 CITARGAATGG CCAAGGAAGA CACTUTUTUTGG CCAACGAAGGA TITCAGAAGCA 240 CCTGAATGGG GGGATAGTA CACGTCCCTC TECATCAGAA AGGCATGTA CCCAACCTCAT 300 ACCCAGACG GACCAGACGAACCATCAT 300 ACCCAGACGACC TITCAGAAAA GUTGTCACCAT 300 ACCCAGACGACC TITCAGAAAAACGAACCAGACC TITCAGAAAAACGAACCAGACC TITCAGAAAAACGAACCAGACCAGACCA CCCTATGAAA GUTGTCACCAT 300 ACCCAGACGACCA CCCTATGAAA GUTGTCACCAT 300 ACCCAGACGACCA CCCTATGAAA GUTGTCACCAT 300 ACCCAGACGACA CCCTATGAAA GUTGTCACCAT 300 ACCCAGACGAC ACCCAGACGAACAACCAGACCAACCAGACCAACAACAACA	WESTERNICH TRECETERE TWOOTERA CACATCACCT CCTCTCCACC	GTCACATCCA	120
CCTGRATGG CCARGACAG TCTCTCTCGG CTACACTAGG CCAGACAGGG TTCAGAAGCA CCTGRATGCG GGGATAGTOA CACGTCCCTC TGCATCAGA AAGGCATGTA CCCAACTCAT ACAGCAGACC TGAGAAACTT CTCTTTCCAA TCC  (2) INFORMATION FOR SEQ ID NO:35: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SINSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 20 of HIP1 (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35: GGCCTGCCA GAAGGTAGA ACGCCCAAGG ACACTCTCT TUGGCTAGTC ATCGCCAGAC AGGGTTCAGA AGCACCTCAA TGCGCGAGA ACTCTCTTC TUGGCTAGTC AAGAAAGGUA TGTAGGCAAC TCATACAAGA AAGGCATGTA CCCAACTCAT AAAACGGGAG GAAGAGGTAT GAAAGGTACA CCATCAACA GACCTCAGAA ACTTCTCTT CCAATCCTGC CACACATCAG TGGACTTCTC CATTCCATAA CCCTGCTGGC CACCTTGACC TGCGTCTAGC TGGCTACTAC TGCCTAAGAG CCCGCACTGAC CACCCACACC TGGCTTCTCC CATTCCATAA CCCTGCTGGC CACCTCACC TGGCATCTCC CATTCCATAA CCCTGCTGGC CACCTTGACC ACCCACGCCA TTGCTCTATGG TTGCTACTACC TGCCTACAGG CCCCACTTGAC CACCCACTCAG TGGTTCACAC TGCCTACAGG CCCCACTTGAC ACCCACGCCA TTGCTCTATGG TGCTTACTCACC TGCCTACAGG CCCCACTTGAC CACCCACTCAG TGGTTCACAC TGCCTACAGG CCCCCACTTGAC CACCCATCAGC TGCTTACTCAC TGCCTACAGG CCCCACTTGAC TGCGACCTCTACCACCACCACCACCACCACCACCACCACCACCA	TOTOTORICE CANCIACCAN CTGGAGAAAA GCTGGAGUUA GTACCTGCCC	TYXXXXXAAAAG	180
ACAGGAAGG CATCTAGGCA ACTCATARAA GGGGAGGAGA CCCTTATGAAA GTGTCACCAT CAACCAGACC TGAGAAACTT CTCTTCCAA TCC  (2) INFORMATION FOR SEQ ID NO:35: (3) SEQUENCE CHARACTERISTICS: (4) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (v) ANTI-SENSE; no (vi) ORIGINAL SOURCE: (A) ORGANISM: builden (x) FEATURE: exon 20 of HIP1 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35: GGCCTGCCCA GAAGGTAAGA ACCCCCAGG ACAGTCCTC TUGGCTAGTC AAGAAAGSCA TGTAGGCAAC TCATACAGA AGCCCTAGAA CCCTACCATCA AAAACGGGAG GAGAGCCTAT GAAGGTGCA CCATCCATAA CCCTGCTGGC CCACTTGACC ACCACATCAG TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCACATCAG TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCACATCAG TGGACTTCTC TGCCTCAGAG CCCCACUTGA CCCTGCCGAC TGGACCATCAG TGGCTACCACC TGCCTCAGAG CCCCACUTGA CCCTGCCGAC TGGACCATCAG TGGACTTCTC TGCCTCAGAG CCCCACUTGA CCCTGCCGAC TGGACCATCAG TGGCTACCACC TGCCTCAGAG CCCCACUTGA CCCTGCCGAC TGGACCATCAG GAGGCCATTGACC TGCCTCAGAG CCCCACCATCAG GAGGCCATTGACC TGCCTCAGAG CCCCACCATCAG GAGGCCATTGACC ACCCACCATCAG GAGGCCATCAC TGGCCCACCACCACCAC GACCCCACCACCAC TGGACCACACCA GACCTCACACACACCACCACACCA	PROBLEM OF THE UNITARIES OF THE PARTICULAR CONTROL OF THE PARTICULAR OF THE PARTICUL	; TTCAGAAGCA	240
(2) INFORMATION FOR SEQ ID NO:35: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 20 of HIP1 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35: OGCCTGCCA GAAGGTAAGA ANGCCAAGG ACACTCCTT TUGGCTACTC AAGAAAGGUA TOTAGGCAAC TOATACAAGA ARGCCATCAT AAAACGGGAG GAGACGTAT ORALGTTCA CCATTCCATAA CCCTGCTGGC CCACTCTGGCAC TIGGTCATGG TEGRATICE CATTCCATAA CCCTGCTGGC CACCTCTGGCAC TIGGTCATGG TEGRATICE CATTCCATAA CCCTGCTGGC CACCTCTGGCAC TIGGTCATGG TOCKTCATCAC TIGGTCATGA CCCTGCCGAC TIGGTCATGG TOCKTCACCAC TIGGTCATGAG CCCTGCCGAC TIGGTCATGG 300 TECRTACACC TIGGTCATGAG CCCCGCCGAC TIGGTCATGG 360	CCTGAATGCG GCGATAGTGA CAGGTCCCTC TGCATCAAGA AAGGCATGTT	ניתנדניטתאקנקאט ב רבייטבייטביטרביט	. 360 . 360
(2) INFORMATION FOR SEQ ID NO:35: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 421 (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 20 of HIPI (xi) SRQUENCE DESCRIPTION: SEQ ID NO: 35: OGCCTGCCA GAAGGTAAGA ARGCCAAGG ACACTCAT TUGGCTAGTC ATCGCAAGAC AGGGTTCAGA AGGACCTCAA TECGGGGAAA GTGACAGGTC CTCTTGCATC AAAAAGGUA 120 TETAGGCAAC TCATACAAGA ARGCCAACAA ACTTCTCTTT CCAATCCTGC ATCGCATCAA 1HIII OAAAGGTTCA CCATCCATAA CCCTGCTGGC CACTCATCAG 300 TECCACCACC TGCCTCAGAG CCCCACCTGA CCCCACCTCAG 300 TECCACCACC TGCCTCAGAG CCCCACCTGAC TGCTCATCG 300 TECCACCACC TGCCTCAGAG CCCCACCTGAC TGCTCATCG 360	ACARGAAGG CATCTAGGCA ACTCATARAR CGGGAGGAGAGA GCATTATORIO	· Gldichern	393
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 421  (B) TYPE: nucleic acid  (C) STRANDEDNESS: double  (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: genomic DNA  (iii) HYPOTHETICAL: no  (iv) ANTI-SENSE: no  (vi) ORIGINAL SOURCE:  (A) ORGANISM: burden  (x) FEATURE: exon 20 of HIP1  (xi) SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GECCTECCE GAAGETRAGA ALCCCAAGE ACACTECTE TUGGCTAGTE ATCGCCAGAC  AGGETTCAGA AGCACTOAA TECEGGATA GIGACAGGTE COTTECATE AAAAAGGUA  TETAGGCAAC TCATACAAGA AAGGUATGTA GCCAACTEAT AAAACGGAG GAGAGGTAT  ING  TOPIAGGTAC CATTCATAA CCCTGCTGGC CACCTTGACC TTGCTCATEG  300  TOCIACCAC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGCTCATGG  360	CAACCAGACC TGAGAAACTT CTCTTTCCAA 15.0		
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 421  (B) TYPE: nucleic acid  (C) STRANDEDNESS: double  (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: genomic DNA  (iii) HYPOTHETICAL: no  (iv) ANTI-SENSE: no  (vi) ORIGINAL SOURCE:  (A) ORGANISM: burden  (x) FEATURE: exon 20 of HIP1  (xi) SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GECCTECCE GAAGETRAGA ALCCCAAGE ACACTECTE TUGGCTAGTE ATCGCCAGAC  AGGETTCAGA AGCACTOAA TECEGGATA GIGACAGGTE COTTECATE AAAAAGGUA  TETAGGCAAC TCATACAAGA AAGGUATGTA GCCAACTEAT AAAACGGAG GAGAGGTAT  ING  TOPIAGGTAC CATTCATAA CCCTGCTGGC CACCTTGACC TTGCTCATEG  300  TOCIACCAC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGCTCATGG  360	(2) INFORMATION FOR SEQ ID NO:35:		
(A) LENGTH: 421  (B) TYPE: nucleic acid  (C) STRANDEDNESS: double  (D) TOPOLOGY: linear  (ii)MOLECULE TYPE: genomic DNA  (iii) HYPOTHETICAL: no  (iv) ANTI-SENSE; no  (vi) ORIGINAL SOURCE:  (A) ORGANISM: human  (x) FEATURE: exon 20 of HIP1  (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  OCCUPACIONAL AGGREGATA ANGECCAAGG ACAGTECTO TUGGUTAGTO ATCGCCAGAC AGGRETICAGA AGGRET	(i) SEQUENCE CHARACTERISTICS:		
(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii)MOLECULE TYPE; genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE; exon 20 of HIPI (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35: (GCCTGCCAG GAAGGTAAGA ANGCCAAGG ACAGTCTCTG TUGGCTAGTG ATGGCCAGAC AGGGTTCAGA AGCACCTCAA TECGGGATA GTGACAGTC CCTCTGCATC AAGAAAGGCA TGTAGGCAAC TCTAGGCAAC AAGAAAGGCA TGTAGGCAAC TCTAGGCAAC AAGAAAGGCA THAGGCAAC TCTAGACCAGAC CCACTTGACC CCACTCAGAC AAGAAAGGCA THAGGGAAC TAGGCAACCAGAC TAGGCCAGAC TAGACAGAC TAGACAGACAGAC TAGACAGAC TAGACAGAC TAGACAGAC TAGACAGAC TAGACAGACAGAC TAGACAGAC TAGACAGAC TAGACAGACAGAC TAGACAGACAGAC TAGACAGACAGACAGACAGACAGAC TAGACAGACAGACAGACAGACAGACAGACAGACAGACAGA			
(C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 20 of HIP1 (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35: GEOCTECCA GAAGSTAGGA ACACCTCAG ACACCTCATC TUGGCTAGTC ATCGCCAGAC AGGGTTCAGA AGCACCTGAA TECGGGGATA GIVACAGGTC CCTCTGCATC AAGAAAGGUA TGTAGGCAAC TCATACAAGA AAGGUATGTA CCCAACCTCAT AAAACGGGAG GAGAGGCTAT CAAAGTGTCA CCATCAACAA GACCTGAGAA ACTTCTCTT CCAATCCTGG CACACATCAG TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTUATGG TGCCACCACCTCACACC GCCCACCTGA CCCTGCCGAC TGTGAGGAC TGGCCATGAC TGTGAGGACC TGCCTCAGAG CCCTGCCGAC TGTGAGGAC TGGCCATGAC TGTGAGGACC TGCCTCAGAG CCCTGCCGAC TGTGAGGAC TGGCTCAGAC TGGCCATGAC	• •		
(ii) MOLECULE TYPE: genomic DNA  (iii) HYPOTHETICAL: no  (iv) ANTI-SENSE: no  (vi) ORIGINAL SOURCE:  (A) ORGANISM: human  (x) FEATURE: exon 20 of HiP1  (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GEOCTECICA GAAGGTAGGA APGECIAGG ACAGTICTE TUGGUTAGTE ATGECIAGAC AGGACTICAT AAGAAAGGUA  TETAGGGAAC TOATACAAGA AAGGCATCAT AAAACGGGAG GAGACGTAT THE CAAAGTGTCA COATCATAA ACCTGAGAA ACTTCTTT CCAATCCTG CACACTCAG 240  TEGACTICTC CATTCCATAA CCCTGCTGGC CCACTTGACC TUGGUTAGTG TOCCACTCAG 300  TECTTACCACC TECCTCAGAG CCCCACCTGA CCCTGCCGAC TUGGUTAGAC 360			
(ii) MOLECULE TYPE: genomic DNA  (iii) HYPOTHETICAL: no  (iv) ANTI-SENSE: no  (vi) ORIGINAL SOURCE:  (A) ORGANISM: human  (x) FEATURE: exon 20 of HIPI  (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GEOCTEUCIA GAAGGTAAGA ANGECCAAGG ACAGTCTCTG TEGGETAGTG ATCGCCAGAC AGGGTTCAGA AGCACCTCAA TEGGEGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGUA TGTAGGCAAC TCATACAAGA AAGGCATCAT AAAACGGAG GAGAGGCTAT CAAAGTGTCA CCATCAACCA GACCTCACAA ACTTCTCTT CCAATCCTGC CACACATCAG TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG 300 TGCCTACCACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGATCATGG 360			
(iii) HYPOTHETICAL: no (iv) ANTI-SENSE; no (vi) ORIGINAL SOURCE: (A) ORGANISM: human (x) FEATURE: exon 20 of HIP1 (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 35: GEOCTECCA GAAGGTAGA APGECAAGG ACAGTCTCTG TUGGUTAGTG ATGGCAGAC 60 AGGGTTCAGA AGCACCTCAA TECEGGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGUA 120 TGTAGGCAAC TCATACAAGA AAGGCATCTA AAAACGGAG GAGAGGCTAT 1HG CAAAGTGTCA CCATCAACA GACCTCACAA ACTTCTCTTT CCAATCCTGC CACACATCAG 240 TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG 300 TCCTTACCACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGGTGACC 360			
(iv) ANTI-SENSE; on  (vi) ORIGINAL SOURCE:  (A) ORGANISM: human  (x) FEATURE: exon 20 of HIP1  (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GGCCTGCCCA GAAGGTAAGA APGCCCAAGG ACAGTCTCTG TCGGCTAGTG ATGGCCAGAC 60  AGGGTTCAGA AGCACCTCAA TGCGGGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGCA 120  TGTAGGCAAC TCATACAAGA AAGGCATGTA CCCAACTCAT AAAACGGGAG GAGAGGCTAT 1HU  CAAGGTGTCA CCATCAACCA GACCTCAGAA ACTTCTCTTT CCAATCCTGC CACACATCAG 240  TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG 300  TGCTTAGGCC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGGTGATGAC 360	<b>7</b> :		
(vi) ORIGINAL SOURCE:  (A) ORGANISM: bumber  (x) FEATURE: exon 20 of HIP1  (xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GGCCTGCCCA GAAGGTAAGA APGCCCAAGG ACAGTCTCTG TUGGCTAGTG ATCGCCAGAC  AGGGTTCAGA AGCACCTCAA TECGGGGATA GTUACAGGTC CCTCTGCATC AAGAAAGGUA  TGTAGGCAAC TCATACAAGA AAGGCATGTA GCCAACTCAT AAAACGGGAG GAGAGGGTAT  CAAGGTGTCA CCATCAACCA GACCTGACAA ACTTCTCTTT CCAATCCTGC CACACATCAG  TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG  TGCTTAGGACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGAGTACC GGGGCATGAC  360	` ·		
(A) ORGANISM: human  (X) FEATURE: exon 20 of HiP1  (Xi)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GGCCTGCCCA GAAGGTAAGA APGCCCAAGG ACAGTCTCTG TUGGCTAGTG ATGGCCAGAC 60  AGGGTTCAGA AGCACCTCAA TECGGGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGUA 120  TGTAGGCAAC TCATACAAGA AAGGCATGTA CCCAACTCAT AAAACGGGAG GAGAGGCTAT 1H0  CAAGTGTCA CCATCAACCA GACCTCAGAA ACTTCTCTTT CCAATCCTGC CACACATCAG 240  TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG 300  TGCTTAGGACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGAGTACC 360			
(x) FEATURE: exon 20 of HIP1  (xi) SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GGCCTGCCCA GAAGGTAAGA ANGGCCAAGG ACAGTCTCTG TUGGCTAGTG ATCGCCAGAC 60  AGGGTTCAGA AGCACCTCAA TECGGGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGCA 120  TGTAGGCAAC TCATACAAGA AAGGCATCTA CCCAACTCAT AAAACGGGAG GAGACGCTAT 100  CAAAGTGTCA CCATCAACCA GACCTCAGAA ACTTCTCTTT CCAATCCTGC CACACATCAG 240  TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG 300  TCCTTAGGACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGAGTACC GGGGCATGAC 360	<b>*</b> '		
(XI)SRQUENCE DESCRIPTION: SEQ ID NO: 35:  GEOTECCEA GAAGGTAGGA APGECCAAGG ACAGTCTCTG TUGGUTAGTG ATGGCCAGAC 60  AGGGTTCAGA AGCACCTCAA TECEGGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGUA 120  TGTAGGCAAC TCATACAAGA AAGGCATGTA CCCAACTCAT AAAACGGGAG GAGAGGCTAT 1H0  CAAAGTGTCA CCATCAACCA GACCTGACAA ACTTCTCTTT CCAATCCTGC CACACATCAG 240  TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG 300  TGCTTAGGACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGAGTACC GGGGCATGAC 360	` ·		
GGCCTGCCCA GAAGGTAAGA ACGCCCAAGG ACAGTCTCTG TUGGCTAGTG ACGCCAGAC 50  AGGCTTCAGA AGCACCTCAA TECGGGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGCA 120  TGTACGCAAC TCATACAAGA AAGGCATCTA CCCAACTCAT AAAACGGGAG GAGAGGCTAT 180  CAAAGTGTCA CCATCAACCA GACCTGACAA ACTTCTCTTT CCAATCCTGC CACACATCAGG 240  TGGACTCCC CATCCATAA CCCTGCTGGC CCACCTGACC ACCCACGCCA TTGCTCATGG 300  TCCCTACCACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGAGTACT GGGGCCATGAC 360			
AGGSTTCAGA AGCACCTGAA TSCEGGGATA GTGACAGGTC CCTCTGCATC AAGAAAGGCA TGTAGGCAAC TCATACAAGA AAGGCATGTA CCCAACTCAT AAAACGGGAG GAGAGGCTAT CAAAGTGTCA CCATCAACCA GACCTGAGAA ACTTCTCTTT CCAATCCTGC CACACATCAG TGGACTTCTC CATTCCATAA CCCTGCTGGC CCACTTGACC ACCCACGCCA TTGCTCATGG TCCTTACCACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGAGTACT GGGGCCATGAC 360	AND ROUGHUE DESCRIPTION, SEQ ID TWO, SO:	е атесселов	c 60
TOTACCCANC TENTACNAGA ARGULATURA CCCNNCTENT ANNACOGUAG GAGACCCTAT 180  CANACTOTO CENTURACUA GACCTORONA RETTETETT CURATECTOC CNENETLATUG 240  TEGROTTOTO CATTOCATRA CCCTGCTGGC CURCTTGACC ACCCNCGCCA TTGCTCATUG 300  TCCCTACCACC TGCCTCNGAG CCCCACCTGA CCCTGCCGAC TGTGAGTACT GGGGCCATGAC 360	PATEMENT OF SOACABIED AT AGD BOTH AACTOOM AS A CAPENDOM	ic <b>aagaaa</b>	W 170
CARLETGICA CCATCAACCA GACCIGACAA ACTICITIT CCAAFCCIGG CACACATCAG 200 TGGACITCIC CATTCOATAA CCCIGCIGGC CCACTIGACC ACCCACGCAA TIGCICATGG 300 TGCTTACACC TGCCICAGAG CCCCACCTGA CCCIGCCGAC TGIGAGTACT GGGGCATGAC 360	PERCENTANCES ATELESCENCE ARACACACTOR TANDESCOUNT TO THE TELESCOUNT	G GAGAGGGTA	ጥ 380
TEGRACTICIC CATTCOATAA CCCTECTEGC CCACTTGACC ACCCACECCA TEGETCATEG 300 TECTIAGRAC TEGETCAGAE CCCCACCTGA CCCTECCGAC TETERETACT GGGCCATGAC 360	PROGRADO TETETETETA CARROLLAS ACCEPTANA ACTETETE TO ACCEPTANCE	;С САСИСИТСА	G 240
TGCCACCACC TGCCTCAGAG CCCCACCIGA CCCTGCCGAC TGTGAGTACT GGGGCATGAC 420	PERIODICA DEPENDACIONED CONTRACTOR OF THE PROPERTY OF THE PROP	A TIGCTUATE	G 300
· · · · · · · · · · · · · · · · · · ·	TGCCACCACC TGCCTCAGAG CCCCACCTGA CCCTGCCGAC TGTGAGTAC	K GCCGGGCGC	A 420

GGGCTGTTCA TEGACCAGGG GAGCAXXXXX CCTTTAAAAG TCTCTGTTCX GCCGGGCGCA

G

420 421

PCT/OS99/11743

## (2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENOTH: 498

WO 99/60986

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS; doubte
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 2) of HiP1.
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 36:

AGGCCCAAGGC	AGGAGAATUU	CTTGAACTCA	CGAGCCCCCAG	TTTGCAGTGA	GCCGAGATGG	60
CGCCACTGCA	CYCCACCCTG	GGCAACAAGA	GCGAGACTCC	ATUTUAAAAA	AAAACTCTCT	120
ATTGUUTTET	ATCTECAGCA	CTGACCGAGO	CCFCTNACCN	GTATGGCAGG	GAAACCCTCG	780
CCTACCTGGC	CUCCCTCCAR	GAAGAGGGAA	GCCTTGAGAA	TGCCGACAGC	ACAGOCATOA	240
GGAACTOCCT	GAGCAAGATC	AAGGCCATCC	CCGNCCTACT	TGGAGTAGTA	TCATTGAGGA	300
GCATTGTTAT	TOTTOTOGGT	GTGCGTGCTG	GTGAATGGCC	ACCCANTUCC	TGATGTTCTC	36C
AGCTAGTICT	TTCTGCACTT	AGARCTTGAT	TOTACIAAAGN	CATTGTTAAA	<b>ATTGGAAAAT</b>	420
CTCCCCGGGT	GCAGTGATTT	<b>አጥርርርጥያ ተለአ</b>	TECCAGEACT	TTGGGAGGCC	GAGTCAGGAG	480
GATCACTIGA	GGCTAGAC					498

- (2) INFORMATION FOR SEQ ID NO:37:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 427
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS; double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: по
- (iv) ANTI-SENSE; no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 22 of HIP!
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 37:

CCCTGTGGGT	TGCAGAAGGT	GTTTGCTGGG	TOCCOCTOCTG	CCTTGCCATU	TIGITAACCOT	60
TACAGATOGC	AGAGGAGAÃG	MANCAGGAGG	CCCCAAGGTO	ACTICACCCT	TTGTGATGTG	120
TTCACAGGAG	CTCCTGCCCA	GGGGAUTGGA	CATCAAGCAC	GAGGAGCTGG	GGGACCIGGT	180
GGACAAGGAG	ATGGCGGCCA	CTTCAGCTCC	TATTGAAACT	GCCACGGCCA	CAATACAGGT	240
AGGAGETTCC	TGCAGGATCT	ССТБЛАЛССА	TGCCTTTGCA	GCTGCCCTTC	TGCAACACTG	300
CTCATTAAAC	ATCTCACACT	CGTTCATTAA	GGCCATGGCA	ACCCCCTAAG	ACAGAAAUCA	360
GAATTTGCCA	GGCACAGTGG	CTUATGOOTO	TANCCCCAGO	ACCTTGGGAG	GATCACTTGA	420
CTCCAGG						427

- (2) INPORMATION FOR SEQ ID NO:38:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTIJ: 367

- (B) TYPB: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE; no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 23 of HIP1
- COSPOUENCE DESCRIPTION: SEG ID NO: 38:

A		-				
CCCCCTGAAT	ACCTTAGAGT	CTGGATTCTT	TTCTGACTCT	CTCAAGAATG	TGGGCAGGGA	бΩ
CTYCCXXXACT	TCCAGATTCA	GGTTTCCCAG	$CTACC\lambda C\lambda CC$	$\boldsymbol{\mathtt{ATGTTGGACT}}$	GAJAGTATAG	120
TAAGACATTA	GYCCATCCTT	AATATTCAAG	GCACATITAG	AAACCATGCT	DATITITIOE	0.80
ACCAGATECT	CAGCAAATCC	CGAGCAGGAG	ACACACCACT	CANNTTGGNG	CTGAATGAAA	240
GGTCGGTCTC	AGCGGCATCC	TGGGACCTAG	GGGAGCAGGA	TCTGTCTTCC	TGACATIGGT	300
CTATACTTTC	CATACTTATT	AGGGAATTAG	AGGAGAGCAG	TACCACCCAC	COCCAACCCC	360
TGAGTTG						367

- (2) INFORMATION FOR SEQ ID NO:39:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 502
- (B) TYPE; nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: homan
- (x) FBATURE: exon 24 of HIP1
- GOSEQUENCE DESCRIPTION: SEO ID NO: 39:

COTCAGORO COTTOTTACO TOCOTTTCCO ATTOCAAGOT 50
AACAGAACC GTCGTGTTCA TTGATCTTGG ATCTTGATCT 120
GTTTGGCAG GATCCTTGCT TCCTGTACCA GCCTCATGCA 180
UGOCTOTAN GGACCTOTAS AGAGAGATTS TOGAGAGOGG 240
COOTGGGCA GGAAGAGGAG GCATCGGTGA CACACTCCCG 300
TOCCOTOTY ACTOMOTOTO TOCACOTGAG TACAGAGCAG 360
GAGGCCCTC CCCIMULGIC AGAGCTCCAG GACCTCCCCA 420
<b>'T</b> 502
CONSTRUCTION OF THE STATE TO SAFETY OF THE S

- (2) INFORMATION FOR SEQ ID NO:40:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 437
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS; double
- (D) TOPOLOGY: linear
- (ii)MOLECULF, TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (ίν) ANTI-SENSE: πο

wn	authaurs.	

- (A) ORGANISM: human
- (x) FEATURE: exon 25 of HIP1

Acriber Lance						
TTTTGGTCTC	TGAATCTTCT	TUTTTTTTTTT	AAAATUGGAA	TACTARTGUT	TATGTUTUAG	60
AGTTACTATG	AGGATCATTT	CCCATAATAT	ΑΓΚΥΤΆΤΑλΑ	OCACCTGCCA	<b>ምእ</b> ምእርያቸለር እጥ	120
CCTCAATAAA	AGGTGECTAT	TACTATTTTT	TATTTCCCTA	GGGTACAGCA	TOCCUTARAG	180
AGTT1TA1GC	CAAGAACITCIT	COMTGGACAG	$\lambda \lambda C C \lambda C T T \lambda T$	${\tt CTCACCCTCC}$	AAGGCTGTGG	240
CCTGGGGAGC	CACTGTCATG	GTGTAAGTAT	CTATTGGTAC	CARGGGTCUT	CCCATGACCC	300
CTCTTCCATT	GATCCACTCC	Αλλελληλος	TARCCACCCA	እእጻለልለስአጥር	TGTCCCTTAG	360
дадтаалста	TTGATCAGGA	AGTUAATAGG	ACCGAGTTTA	CAAGGGAGCC	TGGCTCTCCC	420
AGGGGACACA	CXXXCAGG					437

## (2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 351
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii)MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (jv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 26 of HIP!

## (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 41:

COCACCTEG	CTCTCCCAGG	GGACACAGGG	CAGGGAGGGT	CCCCTCCCTG	TTTAGCCARG	60
GGEGATGGGG	TUGTCTGGAG	GYXXXXTTGT	GGAGGAGTTU	CAGCICATIT	CCCCCTAACC	120
					AGGCAGAGGG	180
					CUAGCITICIC	240
					<u> ምርርምዋ</u> አርአሪያ	300
		GTACTTGGCT				351

### (2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 418
- (B) TYPB: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: human
- (x) FEATURE: exon 27 of HIP1
- (xi)SEQUENCE DESCRIPTION: SEQ ID NO: 42:

CTTTTTATAT GATAGATATG	TUAGGAGUIG	አር <del>ተ</del> እዋእርተርአ	GCAGATTTTG	AGRAGCTOAT	60
TGGTGATTGC CGTTTGGCCC					
<b>ይ</b> ልርተርፈላቸውን፣ ምርርባቸውንልርናሪቸ					180

WO 99/60986 PCT/US99/11743

WO 99/60986	PCT/US99/11743
GCTGATAAGG ACAGCCCCAA CCTAGCCCAG CTGCAGGAGG CCTCTCGGGG AGTG	
RECACTGOOG GOGTTOTGGO CTENACONTO TOCCAMANAT CACACATURA MONG.	
AGCOTTICOA AMAGGACCOT TITOTTACCO ACCOTOTTOA GOTOTTOTOT GUATA	
CTGTGATCCC AAUUAAAYOO CACAGGACTG TGTCTAAATT CTTTCATATT TTTC.	<b>አፓር</b> ፓ 418
(2) INFORMATION FOR SBQ ID NO:43:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 279	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(ii)MOLECULB TYPE: genomic DNA	
(iii) HYPOTHETICAL: no	
(jv) ANTI-SENSE; no	
(vi) ORIGINAL SOURCE:	
(A) ORGANISM: human	
(x) FEATURE: exon 28 of HIP1	
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 43:	
TTTCCACAGA GCATTOCCAT TGGCTGUUTU TCAGGTGCCA GTCAGCCAGG CTAG	antitg 60
ATGAGACCTT CTTGTTTCCA TCCTTCCACA CAACATGGAC TTCTCAAGCA TGAC	
ACAGATCAAA CGCCAAGACA TEGATTCTCA GGTTAGGGTG CTAGAGCTAC AAAA	
GUAGAAGGAC COTCAAAAAC TOGGACACCT TCGGAAAAAG CACTACGAGC TTGC TGCTGAGGGC TOGGAAGAAC GTAAGCTGAC TCAAAGGAT	TGGTGT 240 279
TECTGAGGGC TOSGRAGAR: GTMAGCTGAC TCAMAGGAT	21.5
(2) INFORMATION FOR SEQ ID NO:44:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 3715	
(B) TYPE; nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(ii)MOLECULE TYPE: genomic DNA	
• •	
(iii) HYPOTHETICAL: no	
(iv) ANTI-SENSE: no	
(vi) ORIGINAL SOURCE:	
(A) ORGANISM: human	
(x) FBATURE: exon 29 and partial cds of HIP1	
(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 44:	
ARCATARATT ATCATTGTCT TTTAUGRACA GAGGCATCTC CACCTACACT GCAF	agaagig 60 Paypiac 120
GTAACCGAAA AAGAATAGAG CCAAACCAAC ACCCCATATG TCAGTGTAAA TCCT CTATCTCGTG TGTCTTATTT CCCCAGCCAC AGGCCAAATC CTTGGAGTCC CAGG	- 120 - 180 - 180
CACACCACTG CCATTACCCA GTGCCGAGGA CATGCATGAC ACTTCCCAAA GACT	
ATAGCGACAC CONTYCTCTT TEGACCCATE GTUATUTUTE TTOTTTUCC GCCT	
TTACCATCCA GGCTGGCCAG TGUTGCCCAT CAGCAAGCCT AGGTACGAAG AGX	360 STREETS
GXXXXXAGEG CCAUTCAACA CAGAGGACCA ACATUCAGTU UTGCTGAUTA TTIX	andocco 420
ACAACAATGG GTATCCTTAA TAGAGGAGCT GCTTCTTCTT TCTTGACAGC TTGG	
ARGATOTTAT GOOTTTTCTT TICTGTTTTC TTCTCAGTCT TTTCAGTTIC ATC	
CAAACTTGTG AGCATCAGAG CCCTGATGGA TTCCAAACCA GGACACTACC CTGI CACAGTCAGA AGGACGGCAG GAGTGTCCTG GCTCTGAATC CCAAAGCCAT TCTC	NGAYCYG 600 CCCCCCC 660
TTTGGGOAGT CCCATGGATT TCCACTGCTT CTTATGGTGG TTGGTTGGGT TTTT	17GGTT 720
	· · · · · · · · · · · · · · · · · · ·

GCTGAGTOTO CAGGGCCCCC CAACTGTGGT AGCTCCAGCG ATGGTGCTGC CCACCCCTCT

780

940

PCT/0899/11743

						000
		CACACTGACC				900
		TGACACCTTT				960
		CCGGCGGAAA				1020
		GGYYCACACA				1080
		GCTGGTGAGA				1140
		CAAGCGGGC				1200
		$\tt CTGCAGACGT$				1260
		TGACAAACCC				132D
		TOOTAGECTT				1380
		TINGGAGCAAT				1440
		TCACCCTTCA				1500
		agcaaggaga				1560
		CTCCAACTCC				1620
		ACAGGGCACA				1680
		ATGGACCCCA				1740
		CTATACCAAC				1800
		TTGCTAGGTT				1BG0
TICATCACTG	TGAACCAACC	CCCATCTCCC	TARCCCARRE	CCCTCCCCVV	CTCCCTCTCT	1920
GTGCATTTTC	TAAGTYJGGAC	АТЭСАААААА	CTCTCTCCCA	GGACCTCGGA	TGACCATACT	1980
CACACGTGTG	ACCTCCATAC	TUGGITAAGG	AACTATCAGC	እርባአሬ <mark>አ</mark> ልአ <mark>የ</mark> ፓ	GGGCAGTCTT	2040
AATGTTGAAT	GCTCCTTTCT	CCTTAGTATT	TTTTTTGATTC	AMSSCTCAGA	aggaancetu	210C
DSTGCCTTCC	CTGTCCCAGT	TGPGGCAACT	DOTRACCIAAAC	CHCTCTTCTT	GATGCGGGTC	2160
AACATTTCCA	AAAGTGGCTA	CTCCTCACTT	CTAGATCTCA	OCCATTOTAA	CTCATATGTT	3330
		CGGGCACAGT				2280
AGGCTGAGGT	GGTAGGATUA	CCTGAGGTCA	CCAGTTCAAG	ACCAGCCTGT	CCAACATGGT	2340
					CCCTCCCCCT	2400
AATYCCCACCT	ACTCAGGAGG	CTGAGAGAGG	PODADTAGOA	GAACCCCAGA	GGCAGAGGTT	2460
CCAGTGAGCT	GAGATCACCC	CATTGTACTC	CAGCCTGGGC	AACAAGAGCA	AAACTCCCTC	2.520
					TCAAATTAAG	2580
<b>እ</b> ባባናጥርጥ <u>ር</u> ርር	CAGCCGGGCA	CACTOSCOCIA	TGCCTGTAAT	CCCAGAAUTT	ACCOUNABOOK !	2640
					TAGCAAGACC	2700
46CF4CAC	TARARTTCAR	<u> ሕ</u> ስአርአአለልምፕ	AGCCGGGCAT	GATGGTGCAT	GCCTGTAGTC	2760
TCAGCTAC'I'I	GCCCAGCTAA	GGTGGGAGAA	COCAPTICACE	ŢŢĊĊĠŎŎŎĠŦĊ	GAGGCTGCAG	2620
TOAGCCCTGA	TTGTGCCAGT	GUACTCCGGC	CTOXECTGACA	GAGTGAGACC	AÄASTORTBU	2890
					GAUCCATGAT	2940
ייייייייייייייייייייייייייייייייייייי	ACTOCAGOO	GAGTGACAGA	GCCAGACTCC	<mark>გ</mark> ጥርምርጥምንኢስ	AAACAAACAA	3000
ን አልልጥጥልጥና	GAATGATCY	STOTOTAAAA	далавсслен	GAAATGTTTA	AAAACTTCAT	3060
CCACTTACT	TGAGTCATAZ	CGGTTANGAA	AGURUTTAAA	CAGAAGCAGA	ሊጋርሮይአስጥተርል	3120
ርጥም ነገር ስለገል መ	. AGGAAGTAGI	TOTOAGAGGG	CACATAATTA	CTTTCGTAAT	AGCTCAGATT	3180
ACABOOCTE		' ጥለGእርእእ <b>አጸአ</b> ገ	CAAATTGTCC	TATTGTGACT	AAAAAEOPPO '	3240
71.300.1300.121 71.300.300.121	. GAGGTTATTT	ATGACACAC	: ተፕሬርልፕ <b>ፕ</b> ኢኒስ	ACGGGAATEA	CATCTTARAG	3300
ייייטיי מעשעעער ייייטיי מעשעעער	acymysmaacic	: ርተጥርጥአ <u>አ</u> ልጥ	AGTCACTGAG	CATCACTAGI	CACAMGTOTO	3360
SSANDUNUUU ••••••••••••••••••••••••••••••••	· And reconstit	ነ ጥር ልር ነል ወረፋው ነገር	GGACACCAAC	ANANTCATE	CTTAGGATCG	3 <b>42D</b>
ഗരി ഉയും ഉത്തെ യായ സ്ത്രേസ് (	יים מתממב ה מור המתממב ה	: ጥር-ርነ <u>ል</u> ጥርሃኮር፣ጥባ	የ እርተዋእዲማኮሞ	GTAAGACAGI	GCCCTGAGAC	3480
Calcana co co s No Calcana co co s	. എന്തുന്നുന്നു . എന്തുന്നുന്നു	, <b>դզադլար</b> ը, ,	AGAAAGAAG	CCAGAGATUC	TGTCTCACAC	3540
UTUTAUAUAU ADDA ADDA	. ഉയയാലാന് 1177 പ്രസാധ	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	CCCCTCCAT	. ACCTGAGGTO	AGGAGTTCAA	3600
GTTTAATUU	, POSTANTIAN TAMBUM	. 4009 PBCCCCC	: սեժամիլում : 2200100010	AAATACAAAA	ATTAGCCGGC	3660
GAUTAGUUTU	, WAGGGGGDAM. TOLLANGALANG	i igamaccccci i ateccageti	תנונות בעצמורים. המנונות בענים ביי	. ባኒያልርያርነቦ <u></u> ዉርነርን	GAATC	3715
CATGGTGGC	y GGULIUUTWY	. MILCOMBETY		7.033000minto	,	

## **PCT**





## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C07H 21/04, C12Q 1/68, C07K 5/00

(11) International Publication Number:

WO 99/60986

A3 |

(43) International Publication Date:

2 December 1999 (02.12.99)

(21) International Application Number:

PCT/US99/11743

(22) International Filing Date:

27 May 1999 (27.05.99)

(30) Priority Data:

09/085,199

27 May 1998 (27.05.98)

US

(71) Applicants (for all designated States except US): UNIVER-SITY OF BRITISH COLUMBIA [CA/CA]; University Industry Liaison Office, IRC Building – Room 331, 2194 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3 (CA). MERCK FROSST CANADA & CO. [CA/CA]; PO/CP 1005, Pointe Claire-Dorval, Québec H9R 4P8 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): KALCHMAN, Michael [CA/CA]; #1403-900 Yonge Street, Toronto, Ontario M4W 3P5 (CA). HAYDEN, Michael, R. [US/CA]; 4484 West Seventh, Vancouver, British Columbia V6R 1W9 (CA). HACKAM, Abigail [CA/CA]; 1420 West 11th Avenue, Vancouver, British Columbia V6H 1L2 (CA). CHOPRA, Vikramjit [CA/CA]; Suite 210, 2475 Blenheim Street, Vancouver, British Columbia V6K 4N7 (CA). NICHOLSON, Donald, W. [CA/CA]; 18-750 Milton Street, Montréal, Québec H2X 1W4 (CA). VALLAINCOURT, John, P. [CA/CA]; 18022 Amalfi Street, Québec, Québec H9K 1N7

(CA). RASPER, Dita, M. [CA/CA]; Apartment #7, 16203 Pierrefonds Boulevard, Pierrefonds, Québec H9H 4S8 (CA).

(74) Agent: COPPOLA, Joseph, A.; Merck & Co., Inc., P.O. Box 2000, Rahway, NJ 07065-0907 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published

With international search report.

(88) Date of publication of the international search report:

20 April 2000 (20.04.00)

## (54) Title: APOPTOSIS MODULATORS THAT INTERACT WITH THE HUNTINGTON'S DISEASE GENE

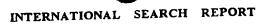
### (57) Abstract

A family of proteins, including a specific human protein designated as HIP1, has been identified that interact differently with the gene product of a normal (16 CAG repeat) and an expanded (>44 CAG repeat) HD gene. Expression of the HIP1 protein was found to be enriched in the brain. Analysis of the sequence of the HIP1 protein indicated that it includes a death effector domain (DED), suggesting an apoptotic function. Thus, it appears that a normal function of Huntingtin may be to bind HIP1 and related apoptosis modulators, reducing its effectiveness in stimulating cell death. Since expanded huntingtin performs this function less well, there is an increase in HIP1-modulated cell death in individuals with an expanded repeat in the HD gene. This understanding of the likely role of huntingtin and HIP1 or related proteins (collectively "HIP-apoptosis modulating proteins") in the pathology of Huntington's disease offers several possibilities for therapy. First, because the function of huntingtin apparently depends at least in part on the ability to interact with HIP-apoptosis modulating proteins, added expression (e.g., via gene therapy) of normal (non-expanded) huntingtin or of the HIP-binding region of huntingtin should provide a therapeutic benefit. Other DED-interacting peptides could also be used to mask and reduce the interaction of HIP-apoptosis modulating proteins with the death signaling complex. Alternatively, a mutant form of HIP-protein from which the DED has been deleted might be introduced, for example using gene therapy techniques. Because HIP-apoptosis modulating proteins have been shown to self-associate, a protein with a deleted DED may compete with endogenous HIP-protein in the formation of these associations, thereby reducing the amount of apoptotically-active HIP-protein.

# FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

	Codes used to identify a		•				
AL AM AT AU AZ BB BB BB BF BG BJ BR CA CF CG CH CI CM CN CU CZ DE DK EE	Albania Armenia Austria Australia Azerbaijan Bosnia and Herzegovina Barbados Belgium Burkina Faso Bulgaria Benin Brazil Belarus Canada Central African Republic Congo Switzerland Côte d'Ivoire Cameroon China Cuba Czech Republic Germany Denmark Estonia	ES FI FR GA GB GE GII GN GR HU IS IT JP KE KG KP LC LI LK LR	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel Iceland Italy Japan Kenya Kyrgyzstan Democratic People's Republic of Korea Republic of Korea Kazakstan Saint Lucia Liechtenstein Sri Lanka Liberia	LS LT LU LV MC MD MG MK ML MN MR MW MX NE NL NO NZ PL PT RO RU SD SE SG	Lesotho Lithuania Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia Mauritania Malawi Mexico Niger Netherlands Norway New Zealand Poland Portugal Romania Russian Federation Sudan Sweden Singapore	SI SK SN SZ TD TG TJ TM TR TT UA UG US UZ VN YU ZW	Slovenia Slovakia Senegal Swaziland Chad Togo Tajikistan Turkey Trinidad and Tobago Ukraine Uganda United States of America Uzbekistan Viet Nam Yugoslavia Zimbabwe



International application No. PCT/US99/11743

A. CLASSIFICATION OF SUBJECT MATTER  IPC(6) :C07H 21/04; C12Q 1/68; C07K 5/00  US CL :536/23.5; 435/6; 530/350  According to International Patent Classification (IPC) or to both national classification and IPC				
B. FIELDS SEARCHED				
Minimum documentation searched (classification system followed by classification symbols)				
U.S. : 536/23.5; 435/6; 530/350				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched				
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) USPAT, CAPLUS, MEDLINE				
C. DOCUMENTS CONSIDERED TO BE RELEVANT				
Category*	Citation of document, with indication, where appropriate, of the relevant passages			Relevant to claim No.
х	WANKER et al., HIP-I: A Huntingtin Interacting Protein Isolated by the Yeast Two-hybrid System. Human Molecular Genetics. March 1997, Vol. 6, No. 3, pages 487-495, see entire document.			1-15
Further documents are listed in the continuation of Box C.  See patent family annex.    later document published after the international filing date or priority				
Special categories of cited documents:      date and not in conflict with the application but cited to under				Discation but cited to understand
1 (	to be of particular relevance of particular relevance;			
	earlier document published on or after the international filing date document which may throw doubts on priority claim(s) or which is	when the de	ocument is taken alone	
	to establish the publication date of another citation or other obsided to establish the publication date of another citation or other considered to involve an inventive combined with one or more other summent referring to an oral disclosure, use, exhibition or other to being obvious to a person skilled in			uch documents, such combination
l l	means document published prior to the international filing date but later than		nember of the same pate	
the priority date claimed		Date of mailing of the international search report		
Date of the actual completion of the international search		3 February 2000 (03.02.00)		
20 DECEMBER 1999  Name and mailing address of the ISA/US		Authorized officer Surrence &		
Commissioner of Patents and Trademarks		SCOTT HOUT	TEMAN	~ /Q
Box PCT Washington, D.C. 20231		Telephone No.	(703) 308-0196	. •

